

# DOCUMENT RESUME

ED 180 786

SE 029 423

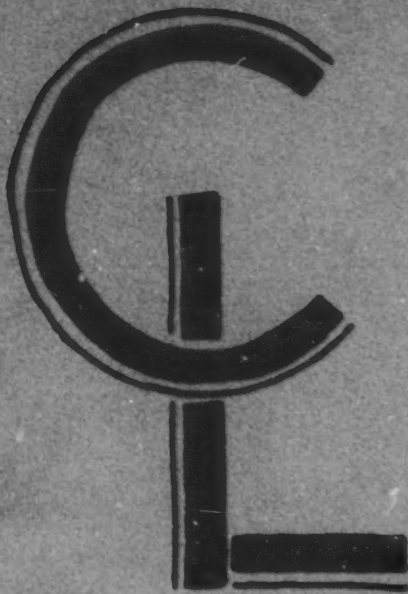
AUTHOR Davis, Robert B.; McKnight, Curtis C.  
 TITLE The Conceptualization of Mathematics Learning as a Foundation of Improved Measurement. Development Report Number 4. Final Report.  
 INSTITUTION Illinois Univ., Urbana. Curriculum Lab.  
 SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C. Basic Skills Group.  
 PUB DATE Oct 79  
 GRANT HEW-NIE-G-76-0085  
 NOTE 105p.  
 EDRS PRICE MF01/PC05 Plus Postage.  
 DESCRIPTORS Behavioral Science Research; \*Behavior Patterns; Classroom Observation Techniques; Computer Assisted Instruction; \*Educational Research; Elementary Secondary Education; Error Patterns; Learning; Mathematical Concepts; \*Mathematics Education; \*Measurement; \*Observation; Test Construction; \*Testing

## ABSTRACT

This is the final report of a research project. The goal of the project was to record a large collection of mathematical performances by students and to develop a system for analyzing these performances. The subjects observed included students in grades 3, 4, 5, 6, 7, and 9 and three adult groups. Four specific results are reported: (1) a large collection of instances of mathematical behavior; (2) a conceptual framework for interpreting and analyzing these instances; (3) a set of specific phenomena of special interest; and (4) some specific suggestions for further work. Also included are chapters entitled The Present Importance of Testing and Education as an Applied Science. (MK)

\*\*\*\*\*  
 \* Reproductions supplied by EDRS are the best that can be made \*  
 \* from the original document. \*  
 \*\*\*\*\*

ED180786



U.S. DEPARTMENT OF HEALTH,  
EDUCATION & WELFARE  
NATIONAL INSTITUTE OF  
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

THE CURRICULUM LABORATORY  
COLLEGE OF EDUCATION  
UNIVERSITY OF ILLINOIS

Development Report  
Number 4

The Conceptualization of Mathematics Learning  
As a Foundation of Improved Measurement

Robert B. Davis  
Curtis C. McKnight

The Curriculum Laboratory  
1210 West Springfield Avenue  
Urbana, Illinois 61801

October 1979

SE 029 423

# TABLE OF CONTENTS

|                                                           | <u>Page</u> |
|-----------------------------------------------------------|-------------|
| I. Why the Obvious Answers are Wrong                      | 1           |
| A. Feedback to Guide Teaching                             | 2           |
| B. Tests to Provide Guidance and Motivation to Students   | 3           |
| C. Other Uses of Tests                                    | 4           |
| D. The Error of Counting                                  | 4           |
| E. What Mathematics Is                                    | 5           |
| F. Criticisms of Present Testing                          | 5           |
| G. Designing Better Tests                                 | 6           |
| II. The Present Importance of Testing                     | 6           |
| III. Education as an Applied Science                      | 8           |
| A. The Task of Description                                | 9           |
| B. Some Factors that Influence Test/Interview Discrepancy | 15          |
| C. The Task of Conceptualization                          | 16          |
| IV. The Goal and Method of the Project                    | 18          |
| A. Observational Methods and Settings                     | 19          |
| B. What to Look For                                       | 23          |
| V. Results                                                | 25          |
| A. The Conceptualization                                  | 25          |
| B. Some Theoretical Questions                             | 30          |
| C. Examples of Behaviors                                  | 35          |
| VI. Corollaries and Implications                          | 58          |
| A. Focused Inquiries                                      | 58          |
| B. Use of the Rhetoric                                    | 76          |
| C. Future Work (Theory)                                   | 84          |
| VII. Educational Science Today                            | 86          |
| A. Points of View                                         | 86          |
| B. Communications                                         | 88          |
| C. Relation to Other Approaches                           | 89          |
| VIII. Details of Specific Activities                      | 93          |
| Bibliography                                              | 96          |

The Conceptualization of Mathematics Learning as a  
Foundation for Improved Measurement

Robert B. Davis and Curtis C. McKnight  
Curriculum Laboratory, University of Illinois

This is the official final report of a research project conducted at the Curriculum Laboratory of the University of Illinois, under the direction of Professor Robert B. Davis, supported by grant number HEW NIE/G-76-0085 from the Basic Skills Group of the National Institute of Education, between the dates of October 1, 1976 and September 30, 1979. The title of the project - "The Conceptualization of Mathematics Learning as a Foundation for Improved Measurement" - is a good brief description. For various reasons one often seeks to measure student achievement (or "knowledge," or "skill") in mathematics; but, in order to measure something, one needs to know quite a bit about it. In the case of mathematics learning and mathematical performance, we do not know very much about "what the thing really is."

*I. Why the Obvious Answers Are Wrong*

The reader who is not a mathematics teacher may be surprised at our caution, because it may seem that there is really no problem. Why not merely write some mathematical questions on a piece of paper, ask students to write out the answers, and then grade each student by counting the per cent of questions that the student answered correctly?

In point of fact, this is probably the most common method of assessing student performance, and it can serve some purposes reasonably well. But we are coming to demand more and more from our testing programs, and consequently we are coming to the point where we must carefully scrutinize the testing programs themselves. This, in turn, leads to a need to look more carefully at what it is

that we are really trying to measure.

Tests are used for many different purposes, and it is necessary to distinguish between some of these different purposes. Consider, for example, just two such purposes:

A. *Feedback to Guide Teaching*

One of the two most common uses of testing is for a teacher to give a test in order to find out what he/she needs to teach next: what misunderstandings need to be corrected? What skills have not yet been mastered? Which students need remedial help, and what sort of help?

This is also probably one of the most successful uses of tests, perhaps because it is a "local" use: the teacher who made up the test also grades it, knows each student personally, can ask oral questions of students to clarify ambiguities, knows the conditions under which the students took the test, and knows what alternative "treatments" are possible for remedial instruction (in the case of failures) and for subsequent instruction on new material (in the case of success).

Yet even at this most local and most informal level, tests are not always adequate. The inadequacies are usually related to the teacher's basic philosophy of teaching. If the teacher sets too low a requirement on mastery, one can get the phenomenon of students "not really learning the material well enough," with the result that they soon forget it, and (in a cumulative subject like mathematics) their future progress is impeded. In the opposite direction, if the teacher sets too high a level of mastery, students can spend very long periods of time polishing a skill which may not require so high a level of polish - with the result that progress is slow, many students may lose interest, and a wrong notion of the irrelevance of school mathematics gets established.<sup>1</sup>

---

<sup>1</sup> One sees this, for example, when some teachers have drilled students to a high level of proficiency in computing in (say) base three numerals, with great emphasis on speed and accuracy.

Problems can also arise if the teacher over-emphasizes "vocabulary drill" and under-emphasizes *performance skills*, so that students are extensively drilled (and tested) on the *words* minuend, subtrahend, abscissa, ordinate, simple closed curve, and so on, but are not able to carry out some of the processes they needed to learn.

Some of these same problems are exacerbated by discontinuities in transitions from one grade, or one school, to another.

#### B. *Tests to Provide Guidance and Motivation to Students*

What is probably the other most common use of tests is to provide guidance and motivation for students. Asked to "learn basic inorganic chemistry" most students would not know *what* to do, and would be disinclined to take the task seriously. But, asked to memorize the law

$$PV = nRT$$

and to be able to use it to solve problems such as:

What volume will be occupied by 0.597 mole of nitrogen at a pressure of 2 atmospheres and a temperature of 27°C?,

these same students may have enough guidance to know just what task they should work on, and if they are told that "the test will be on Friday," they may take the trouble to acquire this skill by that time.

This may well be the most benign use of tests, but even here there are perils. In particular, by choosing tasks and test problems poorly, it is easy to give students a wrong notion of what chemistry is all about. One sees this often in mathematics, where the impression is frequently given that mathematics is entirely a rote subject, involving only the memorization of facts and procedures. When students acquire this expectation, it can stand in the way of their future progress, because the truth is that mathematics often requires that we think for ourselves, sometimes in original and creative ways.

There are other motivational concerns. Suppose, for example, that word problems in an arithmetic test deal only with baseball batting averages. Will this

be fair, given that some student may be deeply interested in baseball, whereas other students are entirely indifferent?

### C. *Other Uses of Tests*

Many other uses of tests are inherently more controversial. Consider, for example, the use of tests to determine who shall be admitted to a particular college. Questions arise as to whether the tests are fair to students from different high schools, or from different cultural backgrounds. Even more fundamentally, it is possible that the tests demand knowledge that is irrelevant to the students' further education, thereby representing gratuitous obstacles that are not justified by educational realities. Or, if the tests deal with obsolete content, they can impede the important task of modernizing the curriculum.

### D. *The Error of Counting*

Nor can we settle for merely counting the number of correct answers, or the percent of correct answers. Consider two computer programs: one may give correct answers most of the time, but the task of revising it so that it will give correct answers *all* of the time may be so difficult as to be unfeasible. The other computer program may never give correct answers, but the flaw in it may be one that is easily detected and easily corrected, after which it *always* gives correct answers. Which is the "better" program?

Precisely similar phenomena occur in math tests. A student who, for the moment, thought that  $5 + 7$  was 13, might use this calculation in three different problems, getting each wrong. Would his performance have been worse if, by chance, the sum  $5 + 7$  had happened to appear in four different problems? Or better, if the sum  $5 + 7$  had happened to occur only once? In any case, however many *answers* the student got wrong, he was in effect making only one single *mistake* - there was but one flaw in his method of dealing with problems. And his one flaw is, hopefully, one that is easily detected and easily corrected. Such a

student may not be far removed from a perfect performance, and merely reporting a count of "three problems with wrong answers" may give a misleading notion of where that student actually stands.

#### E. *What Mathematics Is*

To the lay person it does, indeed, often seem that mathematics is nothing more than a rote subject, dealing with "addition facts" such as

$$3 + 4 = 7,$$

with memorized algorithms (such as the algorithm for "long division"), with *which* of four operations to use in a given word problem, and with not much more than that.

Though commonly held, this idea is an error. It is quite as if we imagined that the subject of "English" was nothing more than memorizing the correct spelling of various words, memorizing the rules for the correct use of punctuation, and memorizing the basic rules of grammar.

What is required in a typical useful "mathematical performance" is as subtle, complex, and creative as what is required in a typical useful piece of writing. Indeed, the two have a great deal in common. Whether one is writing advertising copy, or composing a letter applying for a job, subtle and complex skills - considerably beyond merely spelling and punctuation - are involved. The same is true in using mathematics, whether one is trying to program a computer, design a radio circuit, predict the outcome of an election, or compare the long-term costs of owning two different kinds of automobiles. Hence the testing of entirely rote knowledge of facts and memorized algorithms is not enough to represent the true range of mathematical activities.

#### F. *Criticisms of Present Testing*

Cogent analyses and criticisms of present tests can be found in: Tyler, et al., 1979A; Zacharias, 1979; Davis and Henkin, 1979; Henkin and Davis, 1979; Hilliard, 1979; Padilla, 1979; Schwartz and Taylor, 1979; Krutetskii, 1976; Davis, Romberg, Rachlin, and Kantowski, 1979; Houts, 1977; Zacharias, 1977;

Morrison, 1977; Schwartz, 1977A; Hoffman, 1962; Schwartz, 1977B; Taylor, 1977; and elsewhere.

### G. *Designing Better Tests*

One could imagine an optimized curriculum-and-testing program, where what was learned was almost exactly what *needed* to be learned, where the sequential order and timing were correct, where everything was learned to an appropriate level of mastery, and where skills and knowledge were appropriately maintained by suitable review. One can *imagine* such a curriculum-and-testing program, but that probably isn't what one can see in typical practice today (although much present practice is surprisingly good).

Basic to the present research project is the idea that, in order to design and construct such an optimized program, one needs more knowledge, on a fundamental level, concerning the nature of mathematical performances and mathematics learning. As James Greeno has remarked: if you want to *compare* two radios, your task is simple; you do the sorts of things that *Consumers Union* does - use listening panels, test for sensitivity, noise rejection, etc. If you want to *repair* a radio, you need to have knowledge about how radios work, i.e. about what goes on inside of a radio. But if you want to *design a better radio*, you need to know about some of the basic laws of physics.<sup>2</sup>

Acquiring some of this fundamental knowledge was the goal of the present project.

## II. *The Present Importance of Testing*

Tests can determine who graduates from high school and who does not, who gets into which college, or who gets which job. Tests can be used in evaluating schools and curricula, as well as students. For many reasons, in recent years tests have become far more important than ever before. These reasons include:

---

<sup>2</sup> Paraphrased from remarks made at the N.I.E. Conference on Testing, Teaching, and Learning, August, 1978 [Cf. Tyler, et al., 1979A; 1979B].

- 1) The introduction of new instructional programs, based on novel assumptions: At the present time we see new curricula in many schools (and many new hybrids of earlier curricula). There are a sizeable number of locally-made programs in use, or coming into use. Computer-assisted instruction is, after 15 years of false starts, beginning to make a serious appearance in schools (Philadelphia; Chicago; Los Angeles; Texas; Illinois; and elsewhere). A new "economical" curriculum (based on a novel allocation of resources) is being developed at University High School in Urbana, Illinois. Some of these changes are rather sizeable ones, even radical departures from tradition.

New curricula and new programs open the door to new possibilities, *and also to new risks*. Careful measurement of learning outcomes is essential (and our commonly-used present evaluation procedures are particularly ill-suited to the comparison of different curricula) [CBMS, 1975, pp. 123-4].

- ii) The use of student evaluation *within* programs. Many of these new programs - especially CAI and individualized programs - make an essential use of internal measurement *within* the curriculum in order to route or pace students properly. With 25 students pursuing, because of individualization, 25 different programs or lessons, the teacher cannot be involved in every decision, and the "automatic" pacing and routing become far more important.
- iii) "Objectivity" and "fairness" - present national goals of equal educational opportunity mean, necessarily, greater reliance on "objective" measures of achievement and of readiness, as a basis for program decisions;
- iv) General public doubt about school effectiveness. There is widespread questioning of the educational effectiveness of our schools today, especially in areas such as mathematics. (One specific instance might be the effectiveness of compensatory education programs (Averch et al., 1972, pp. 100-125; Stanley, 1973; DHEW, 1972; Gordon and Wilkerson, 1966; Bissel, 1973). On individualized instruction, cf., e.g., Weber, 1976.) The only appropriate response requires careful measurement of fundamental cognitive growth.

In fact, one can summarize the case for a new look at testing approximately as follows<sup>3</sup>:

1. Tests are being asked to play a larger role - and more different roles - than ever before;
2. The diversity of new curricula that have been developed since the 1950's - and that must continue to be developed - require rapid adjustments to testing new content areas, and testing for new *kinds*

---

<sup>3</sup>

Cf. Mays, Melmed, and Scribner, 1979.

of "skills" and "knowledge."

3. Modern technology - especially computers - create new opportunities for testing, often in quite novel ways;
4. Modern cognitive studies have carried us into new and deeper understandings of the nature of human thought. As George Miller has put the matter:

"The theory that has, explicitly or implicitly, dominated most thinking about tests is that they provide measures, analogous to length and weight, that can be used to arrange children and what they have learned, in a simple linear order. This theory supported initial research on testing, but it is obviously inadequate to our present conceptions of the complex nature of human knowledge." [Miller, 1979]

In all of the important practical matters where tests are used today - from affirmative action programs to more effective CAI - inadequacies in our basic measurement procedures can have immediate consequences, both for individual students and for the shaping of instructional programs.

### III. *Education as Applied Science*

Is educational practice an applied art, or is it an applied science? Most practitioners probably regard it as an applied art, and they are probably not wrong to do so, as Hawkins [1966], among others, has pointed out. But the present project has sought *scientific* knowledge, and it may be well to consider briefly what that means in the case of education.

Given the complexity of educational problems and educational practice, it is no wonder that education has turned to science for guidance [Cremin, 1961; Callahan, 1962]. In the early years of doing this, the major emphasis was on *measurement* and *experiment*. The pay-off from this effort has generally been less than had been hoped for. More recently, a more sophisticated analysis has been made of what "science" has meant in the cases of physics, astronomy, chemistry, biology, medicine, and other "hard" sciences. In these areas, the pay-off from science has been very great, indeed. What has become clear is that, important though

*measurement and experiment* can be in science, a very large share of the effort - indeed, the larger part of it - must go to *description* and to the development of *appropriate conceptualizations*. In the attempts at bringing science to bear upon educational problems and educational practice, far too little effort has been expended on the *description* and *conceptualization* tasks.

[The name of the disease *diabetes mellitus* is itself an intriguing commentary on the vast two-thousand-year effort at *description* that has gone into the development of the physical and biological sciences. "Mellitus" is a cognate of *mellifluous* and *mellite*, from the Latin word *mel*, meaning "honey." Why "honey"? Because, in the course of efforts to identify the disease and to describe it adequately, a physician had *tasted the urine* of typical patients and found it to taste like honey. The successful sciences of today are built on the foundation of vast and heroic efforts at *description*.]

#### A. *The Task of Description*

Suppose that we accept the possibility that the describing of a student's mathematical competences may be a complex, non-trivial task. We can use modern technology to help us study how this student deals with mathematical problems: we can give the student a problem to work on, and while he works on it, and perhaps discusses with us how he is doing it, we can record his performance on audiotape or videotape.

Stanley Erlwanger did this, a few years ago, with some students in grades 4,5, and 6. What he found was disquieting.

The discrepancy between student mathematical behavior as revealed by multiple-choice paper-and-pencil tests, vs. student mathematical behavior revealed by interview tasks, is great and complex. We give some very brief notion of this discrepancy by citing two specific instances:

1. Erlwanger's "Benny". Erlwanger (1973) reports an extended sequence of interviews with "Benny," a 12-year-old 6th grade boy studying in an

individualized curriculum. His studies for 5 years had been guided not by teacher judgment, but rather by "automatic" pacing and sequencing by curriculum-embedded tests. As Erlwanger reports, "he was making much better than average progress" through the individualized curriculum, "his teacher regarded him as one of her best pupils in mathematics," and the teacher assumed "that Benny could not have progressed so far without an adequate understanding and mastery of previous work" (p. 7). The interview data presented rather a different picture. Asked to add  $\frac{2}{1} + \frac{1}{2}$ , Benny wrote  $\frac{2}{1} + \frac{1}{2} = 1$ ; asked to convert  $\frac{2}{10}$  to a decimal, Benny wrote  $\frac{2}{10} = 1.2$  as his answer. The details of Benny's answers are important, and we shall return to them below. But, for the moment, here is one example of the discrepancy between the behavior revealed by multiple-choice tests, and the behavior revealed by interview tasks.

2. Erlwanger's "Mat". Another student whom Erlwanger studied, via an extended sequence of interviews, was an 11-year-old fifth grade boy with an I.Q. score of 121, whom Erlwanger calls "Mat." Mat had been in an "individualized" school math program for grades 2,3,4, and 5, so that his personal curriculum had been determined "automatically," by tests embedded within the curriculum.

Erlwanger reports the initial data on Mat as follows:

"His teacher considered him to be above average in mathematics. She believed that he usually understood the basic skills and concepts in his work...

"Mat was confident about his work. He was prepared to defend his point of view in discussions...

"From the comments of his teacher and his rapid progress in the [school instructional] program it appeared that Mat understood his

work." [Erlwanger, 1974, p. 63]

As usual, the interviews revealed a quite different situation:

"...in the course of discussions and interviews with him, it was discovered that Mat had his own conception of mathematics as a system of methods for finding answers. He believed that 'You get a different answer every [different] method you use,' on a particular problem. Further, the required answer from these methods depended upon the directions he was given. We shall examine various aspects of Mat's ...conception to see how it guided his thinking and how he succeeded in mathematics." [Erlwanger, *ibid*, p. 64].

Mat's actual work in arithmetic can be indicated by the following excerpt from one of the interviews:

"Mat had recently completed level F multiplication in which he learned how to multiply numbers involving more than two digits, by using various methods. According to his test scores, and the comments of the teacher, the prescriber and the grader, Mat had a very good knowledge and understanding of multiplication involving whole numbers.

When Mat was asked to multiply  $36 \times 27$ , he wrote  $\overset{36}{\times 27}$  and said:

"Well, this...*this kind of multiplication I am good at. See, you cover up...this number...the 2 on the 27 first. And then you take the 6 times the 7...then put the answer down there [writing 42 on scratch paper]...that's 42. So put the 2 down there [below the 7] and carry the 4 [above the 3]. And then you take the...you still cover up the 2...and take the 3 times the 7... is 21; but then you add the 4 up here...which will be 25. So you put 25 down here. And then for the 2, you have to add a zero down here.*

Continuing in this way he obtained 972 as his answer. To check his

answer, he said: "The way you check these problems is...you just switch them around... like [writing  $\overset{27}{\times 36}$  ] there, instead of  $36 \times 27$ ." Again he obtained 972.

However, Mat also obtained three different answers for  $200 \times 100$  when the numbers were arranged differently:

$$\begin{array}{r} 200 \\ \times 100 \\ \hline 20000 \end{array} \quad \begin{array}{r} 200 \\ \times 100 \\ \hline 2000 \end{array} \quad \begin{array}{r} 200 \\ \times 100 \\ \hline 200 \end{array}$$

We then had the following discussion:

Excerpt 18:

1. E: Now which is the right answer?
2. M: *They're all right! It's just that the numbers are in different positions.*
3. E: Would it work this way in addition and subtraction...and so forth?
4. M: [pause] Well, let's take 200 plus 100. [Mat wrote  $\overset{200}{+} \overset{200}{100}, + 100$ , and  $\overset{200}{+100}$ ]. *You wouldn't get the same answers now...but...300...O.K.?... and 200...100...there..O.K. now [Mat wrote down his answers as 300, 2100, and 20100 from right to left, respectively].*
5. E: And you say they are all different?
6. M: Uh uh.
7. E: Would they [booklets] use all these three methods in mathematics?
8. M: *Not at the same time...maybe on different pages...ya.*
9. E: Now where would you come across an example like  $200 + 100$ ...like this...with the 1 below the first zero [i.e.,  $\overset{200}{+} \overset{100}{100}$  ] ?
10. M: In addition.
11. E: Have you ever done some like that?
12. M: I think so..ya, ya...I have. It's been in the third grade though... because I was in E addition in third grade.
13. E: Uh uh. Now suppose you...you had to take a test. O.K.? And they had this example down..  $520 + 360$ ...written across [i.e., horizontally].  
How would you do it? [Mat worked as follows:  $\overset{520}{+360}$  ]  
880

14. E: How would you know they wanted you to put it that way?
15. M: Well...another way to do it is to move this [520] down here [Mat  
360  
wrote +520 ].
16. E: Why would you use that one?
17. M: *Because that's the way...that's the way it's usually written.*
18. E: Pardon?
19. M: [somewhat excited] *That's the way they are!* You see, in the booklet...  
they...they...on the first page sometimes they show you...like  
this goes here, this goes here, and this goes here [pointing at  
each example].
20. E: Oh...you mean they...they tell...
21. M: [interrupting] *They show you at the first [page] of the book.  
They don't tell you in the test...but when you get it in the book,  
they tell you how to do it. [Erlwanger, 1974, pp. 86-7]*

Nor is it only the case that students know *less* than one expects. Sometimes careful observation reveals that they know *more*:

3. Kye's Subtraction: This mathematical behavior, by a third grade boy named Kye, was observed during a classroom lesson. The teacher, working on the blackboard, was showing "subtraction with borrowing [or "re-grouping"]," using the problem

$$\begin{array}{r} 64 \\ -28 \\ \hline \end{array}$$

Teacher: You can't take 8 from 4, so...

Kye [interrupting]: Oh, yes you can! [Kye went to the board]:

Kye said:

Four minus eight is negative four...

...and sixty minus twenty is forty...

Kye wrote:

$$\begin{array}{r} 64 \\ -28 \\ \hline -4 \end{array}$$

$$\begin{array}{r} 64 \\ -28 \\ \hline -4 \\ 40 \end{array}$$

...and forty and negative four give you  
thirty six

$$\begin{array}{r} 64 \\ -28 \\ \hline -4 \\ 40 \\ \hline 36 \end{array}$$

[This algorithm was unknown to all of the adults in the situation, and was therefore, in this sense, an original discovery by a third-grade student.] [Cochran, Barson, & Davis, 1970]

4. Ricky's Algebra. Ricky, a seventh grade student, was in a class that had completed a unit on solving equations. The textbook emphasized a pre-ordained "formal" procedure: combine like terms; multiply to eliminate fractions; and so on. The teacher had departed from this, to stress a heuristic approach along the lines of George Polya: What things in the problem are especially hard? Can you eliminate these obstacles? Can you turn the problem into something more familiar?

Ricky came to think of equation-solving in a third way, quite different from both the textbook's formal methods, and the teacher's heuristic methods. (Ricky's approach is sometimes called an "informal" approach, but a better name might be "opportunistic" - take maximum advantage of special features of that particular problem.)

Ricky solved for  $x$  in the equation

$$\frac{5}{x} = \frac{10}{3x - 7}$$

by saying, almost immediately, that  $x$  had to be 7. Here (not quite verbatim) is Ricky's subsequent explanation of how he knew that " $x$  must be 7."

"Well, [going from the left-hand side of the equation to the right-hand side], they've doubled the numerator. [Therefore]...they must have doubled the denominator. But instead of  $2x$ , they put  $3x$ , so there's one  $x$  too many.

But then they fix up the extra  $x$  by subtracting 7. But if taking away 7 gets rid of the extra  $x$ , then  $x$  must be 7." [from a protocol in the collection of the authors].

The road which children travel in learning mathematics looks very different when assessed by interview procedures from the way it looks when assessed by commonly-used standardized tests. Not only are the usual "measurements" wrong, but they are derived from incorrect basic conceptualizations of the nature of learning.

B. *Some Factors that Influence This Discrepancy*

That such a discrepancy can exist is not entirely mysterious. Some causes are:

- 1) In order to achieve maximum standardization, traditional testing has limited attention to a severely restricted list of possible alternative behaviors, usually selecting one out of four or five possible "answers" from a pre-specified list. Thus, Mat would usually have no opportunity to write

$$\begin{array}{r} 100 \\ + 200 \\ \hline \end{array},$$

nor to demonstrate his belief that the problem does not determine a unique correct answer.

- ii) Traditional tests are based upon a narrow "strip" curriculum. Experiences outside of this narrow strip are not provided for. Thus, a third-grader like Kye would usually have no opportunity to demonstrate skill in using negative numbers.
- iii) Multiple-choice tests often omit important wrong answers. Thus, one widely used test (CTBS, intermediate grade level, Form P) includes the question

$$6 \div \frac{1}{2} = ,$$

but omits the most common wrong answer (namely, 3).

- iv) Commonly-used tests are based on a narrow range of behaviors, mainly rote recall of algorithmic procedures and rote recall of statements. *Not* usually included are critical skills ("What's wrong in the following discussion?"), nor improving upon a partial solution to a problem, nor devising heuristic approaches to a problem, nor learning something new, nor discovering patterns, nor recognizing ambiguities - even though these behaviors are even more important for success in learning mathematics.

For these, and other, reasons, presently-used tests do not do a good job of choosing between different curricula, nor in guiding "individualized" curricula for individual students. But the fundamental error is that most present tests do NOT provide us appropriate "peepholes" to look through surface phenomena and discern a bedrock foundation of the kind of cognitive growth that must occur if one is to "learn mathematics."

[A somewhat similar analysis has been given in Krutetskii, 1976.]

To put the matter a different way, one needs a better conceptualization of the fundamental processes of "learning mathematics" and "working on mathematical tasks."

### C. *The Task of Conceptualization*

No task in science is more important than the task of *conceptualizing* aspects of reality. Yet somehow this task has been neglected in education's early attempts to make use of science.

Science deals with reality, but it deals with reality by means of *ideas*. These ideas do NOT come directly from reality. Scientists *invent* them.

When Isaac Newton proclaimed the law that, in the absence of forces, an object would move in a straight line with a constant velocity, what was he doing? *Describing* something that he had observed? By no means - where on earth could Newton have observed an object that was not acted on by any forces? Was he generalizing from a large number of instances? Again, surely not! He had not

one single instance, much less a large number of them.

On the contrary, he was making up the rules of the game: in effect he was saying, if something were to move in a straight line at a constant speed, I would not look for any forces propelling it. I would take that to be its normal behavior. *But if an object moves in any other way, then I must find some arrangement of forces that will explain its motion.*

In a similar way, Lavoisier (as Kuhn, 1962, points out) *invented* the chemical elements: that is to say, he *conjectured* that there were certain chemical substances that were NOT made up of other substances. Call these the "elements." All *other* substances, Lavoisier conjectured, were made up of combinations of various of these elements.<sup>4</sup> That still left the task of fitting reality into Lavoisier's scheme by, among other things, discovering which substances actually were elements... but at least Lavoisier had defined the general nature of the task, and in a quite new way. He had introduced the *idea* of chemical elements vs. compounds.

Similar steps were taken by Mendeleev in proclaiming his Periodic Table. In still another case, the introduction of the *idea* of the nuclear atom was carried out by, essentially, *invention*.

Many of these inventions have a certain *metaphoric* quality: the idea of the nuclear atom, with orbiting electrons, almost certainly owed something to the known configuration of the solar system, with orbiting planets.

Where are we to look for the basic conceptualizations of education - or more specifically, of "thinking mathematically"? This is discussed at some length in Davis and McKnight, 1979A.

Essentially, our answer is that we can best begin by looking at the information-processing activities of digital computers. This answer has also been

---

<sup>4</sup> A point of view on science, very similar to the one given here, has been stated by Herbert Simon (1979).

given by Klahr [1976], Simon [1979], Heller and Greeno [1979], Winston [1978], Newell and Simon [1972], Brown and Burton [1978], Goldstein [1979], John Seely Brown [1979], Matz [1979], Minsky [1975], Minsky and Papert [1972], Bobrow [1975], Bobrow and Norman [1975], Rumelhart and Ortony [1977], Schank [1975], Van Lehn and Brown [1979] and many others. We refer to the discussion in Davis and McKnight, 1979A.

The main ingredients in our conceptualization of mathematical thought are:

- procedure
- sub-procedure
- super-procedure
- unit sub-procedure
- VMS sequence
- integrated sequence
- frames
- assimilation paradigm
- heuristics
- meta-language (or "planning language").

For definitions, we refer to Davis and McKnight, [1979A].

#### IV. *The Goal and Method of the Project*

Learning mathematics can be analyzed as the combination of three components:

- i) an "information processing" or "cognitive" component
- ii) a motivational component
- iii) "environmental" components: teachers, textbooks, school schedules, peer influences, home influences, etc.

This project has intended to focus on the first of these three - the "cognitive" or "information processing" activities - and for the most part has done so, but in a real world with real children it has sometimes been unavoidable to include occasional aspects of the other two.

The goal has been to record a large collection of mathematical performances by students, and to develop a system for analyzing these performances.

Very broadly speaking, this is not an unusual sort of goal. It is, very broadly, what chemistry did in identifying very many chemical substances, observing and recording their behavior in various circumstances - solubility in water, reactions with other chemicals, boiling and melting points, etc. - and inventing conceptual schemes such as the nuclear atom and the Periodic Table to make possible the analysis of these observed behaviors.

We have in fact assembled a sizeable collection of instances of mathematical behavior, and have developed a conceptualization of the relevant cognitive processes that does allow effective analysis of these instances. In this sense, we have done what we set out to do. But it is in the nature of this work that the details are important, and we must consequently consider various details. This, in turn, means that reports are necessarily lengthy. We shall not reproduce instances here, except very briefly, but refer instead to full accounts that appear elsewhere.

[The best *brief* account of this work appears in Davis and McKnight, 1979A. A different aspect of the work is presented in Davis, McKnight, Parker, and Elrick, 1979. The most complete account is Davis, Jockusch, and McKnight, 1978.]

A. *Observational Methods and Settings*

1. The human beings whom we have observed have included:
  - a. students in grades 3,4,5,6,7 and 9, of varied SES, racial, and ethnic backgrounds
  - b. adults studying mathematics at a neighboring community college
  - c. adult "experts," mainly either mathematicians or mathematics teachers.
  - d. adults learning to use hand-held calculators.
2. The settings in which observations have been made have included:
  - a. classroom mathematics lessons

- b. task-based clinical interviews
- c. students working at PLATO CAI terminals
- d. self-reports by adult experts

Each of these requires some explanation.

*Observation of students as they work at PLATO CAI terminals.* We have the good fortune to have PLATO computer terminals installed in various local classrooms, where they are used to deliver instruction in mathematics in a "computer-assisted instruction" teaching mode. These terminals can, and do, administer tests (although they can do this unobtrusively, by merely gathering data on regular student performances). These terminals are at present primarily in grades 3,4, and 5, where they provide instruction in arithmetic and in beginning algebra to students of extremely diverse backgrounds, although we have also had some more limited chances to observe younger children on PLATO terminals. In addition, through the cooperation of Ms. Marge Stillwell and South Side School in Champaign, Illinois, we have been able to observe some remedial students (in a "Title I" program) as they worked on PLATO. In order to test a specific conjecture, Philip Parker programmed a special test that was administered, via PLATO, to 28 students in grade 5. Except for the Parker test, where data was collected only through the computer itself, the procedure is to observe students directly, though mostly unobtrusively, as they work at the computer terminal. In some cases the interviewer interrupts to pose an oral question in order to clarify some particular point (or the interviewer may wait until the end of the session to pose his oral questions).

Observation of students as they work at PLATO CAI terminals has been the least important of the three settings, as far as the present study is concerned, primarily because greater flexibility can be achieved with task-based interviews, especially in terms of probing various content areas. Nonetheless, CAI terminals do have some advantages, and in studies where the relevant content was sufficiently

circumscribed, the observation of students at terminals could well be the most effective approach.

#### *Observation of Classroom Lessons*

In the present study, the observation of classroom lessons has played a somewhat unusual role, for four main reasons:

- 1) Observations have been mainly at University High School, a laboratory school owned and operated by the Curriculum Laboratory of the University of Illinois. This is a school for gifted students. Hence *this portion of the study* deals with a non-typical student population. It is important to note that this applies *only* to this part of the study. The PLATO terminals are in typical public schools, such as Booker T. Washington and South Side schools in Champaign, Illinois, and deal with the full range of typical public school student populations. The Title I students are low SES remedial students, with a heavy minority representation. The community college students were, presumably, not gifted in the area of mathematics.

In fact, the difference is not important. *At the fundamental level of information processing which is the main concern of this study, almost all students exhibit the same phenomena*, whether this be a simple misreading of one's own handwriting that confuses an  $x$  with a  $y$ , or whether it be a student saying "I can't add  $x+3$  because I don't know what  $x$  is", or even a student who makes the sophisticated epistemological mistakes discussed in Davis, Jockusch, and McKnight, 1978.

- 11) The principal investigators in this study have arranged to be the "regular" mathematics teachers for the relevant classes. There are trade-offs in such an arrangement, but it is clearly true that *what an observer sees will tend to be different from what a teacher sees*. For this part of the observational program, we wanted to get the

teacher's view of what was happening in the course of instruction.

- iii) As befits a school for gifted students, the mathematical content at University High School is unusually advanced. (Specimens of student work can be seen in *The Journal of Children's Mathematical Behavior*, vol. 2, no. 2 (Spring, 1979).)
- iv) We teach the same students - except for a small number of transfers - for 5 consecutive years. Hence we can relate their progress in one year to their progress in earlier or later years. (The effect of this is felt less in the present study, since it dealt primarily with younger students.)

#### *Observation Via Task-Based Interviews*

This, in one or another variant, is by far our most common, and most important, method. A student is given a mathematical task, either written on paper, or presented orally by an interviewer. The student has pen or pencil (pen has the advantage of leaving a more permanent record when students change their minds) and a supply of paper. If appropriate, he may also have a protractor, ruler, compasses, a hand-held calculator, etc. There is an adult interviewer, who has his own pad of paper and keeps his own notes. There may also be a second adult present as an additional observer. If present, this second adult may or may not participate in the conversation.

The student's behavior is usually, but not always, recorded on audiotape, and less frequently on videotape. These decisions are often influenced by opportunity.

Interviews may be formally scheduled and occur in a carefully prepared setting, or they may be less formal - some have even occurred when students at athletic events have sought help on mathematics problems. Taking advantage of informal settings can be very valuable.

Students work on the mathematics tasks while the observer(s) watch. Students

are asked to verbalize as fully as possible while they work, and most do so with little apparent difficulty.

There is great variability in the degree of obtrusiveness or unobtrusiveness of the adult observation. At one extreme, when a desired behavior is directly observable, and when it is desired to provide maximum guarantees that student behavior not be influenced by the adults, the interviewer/observer may remain entirely silent and unobtrusive for the whole time that the student works on the task. At the opposite extreme, it is sometimes crucial to determine, as nearly as possible, *exactly how* a student thought about some problem. In this case, one (or both, if two are present) adult(s) may pursue a fairly insistent line of questioning. This can take one of several forms:

- (a) Asking "can you explain that to me?";
- (b) Asking "how did you [or "why did you..."] decide that...?";
- (c) Commenting: "Suppose I told you that I think there is one mistake on paper...Do you suppose you could find it?";
- (d) Posing a *new question* that may reveal something about how the student is thinking - e.g.: "How would you solve this problem if that were an  $x^2$  term instead of an  $x^3$  term?"

Typed transcriptions are subsequently made from audio tapes, usually with the help of the interviewer or observer, since they are better able to interpret some of the less clearly audible (or recognizable) sounds on the tapes.

#### *Self-Reports From Adult Experts*

In a small number of cases we have been able to get written reports from "experts", as they work on some mathematical task, especially one which initially confuses them.

#### B. *What to Look For?*

Of course, setting and "observational procedures" are only part of the story. A more central question is: what is one looking for?

Or even: How does one decide what to look for?

Our method has been deliberately opportunistic. We start with an attempt to observe *everything* that can be observed and that seems to have relevance to how students think about mathematics. Of course one cannot in fact observe "everything," but this beginning has produced a good diversity of various kinds of behaviors (cf. Davis, Jockusch, and McKnight, 1978) that can become a basis for further observation.

The direction of future observations is shaped by four considerations:

- i) deliberately giving preference to matters that relate to the conceptual analysis that is being built up;
- ii) deliberately giving preference to matters that are confusing (to the observers, not to the students) or controversial (in the sense that observers disagree, either about what happened, or about the interpretation of what happened);
- iii) deliberately giving preference to phenomena that re-occur with considerable regularity;
- iv) unavoidably giving preference to phenomena that are conspicuous or easy to observe.

This last category - "conspicuous or easy to observe" - includes: (a) behaviors that lead to errors or wrong answers; (b) behaviors that show great variation from one student to another; and, presumably, (c) other categories which we are not consciously aware of.

To give a clearer idea, then, of the observational method it is necessary to consider what we were looking for, and to clarify this we need to consider the theoretical conceptual structure that emerged. We turn to this next.

It should be noted that the theoretical structure did NOT precede the observations; both were built up simultaneously, and the present state of each influenced the further development of the other.

### V. Results

The results of the project consist of:

- i) a large collection of instances of mathematical behavior
- ii) a conceptual framework for interpreting and analyzing these instances
- iii) a set of specific phenomena of special interest
- iv) some specific suggestions for further work.

#### A. The Conceptualization

Fundamentally, human thinking about mathematics is analyzed by analogies with sophisticated information processing by computers. We do not here refer to the use of a computer to add or subtract, nor even to solve complex mathematical problems of a routine "numerical analysis" type. Rather, we refer to those computer programs that interpret photographs, play checkers or chess, compose music, write poetry, carry on conversations, control mechanical devices, and - indeed - even learn algebra by "reading" a textbook. One need not assume that computers "think" the same way that humans do - our private guess is that they do not. The point is that sophisticated computer programming does seem able to provide useful *concepts* for analyzing information processing in other contexts. We have listed earlier the key concepts in our formalization; we now provide brief descriptions, referring to Davis and McKnight (1979-A) for further discussion.

1. *Procedures.* A *procedure* is a definite sequence of steps, aimed at achieving some quite specific result. (Think of it as a human information-processing analogue of a *program* in a programmable hand-held calculator.) A procedure can relate to other procedures in several ways, the most important being that Procedure A can call upon Procedure B, in order to accomplish part of the assignment of Procedure A. In such a case,

Procedure B is said to be a *sub-procedure* of Procedure A, and A is said to be a *super-procedure*. A well-known result from Erlwanger's studies of 4th, 5th, and 6th grade arithmetic students is that their arithmetic errors are nearly always located in super-procedures; the various sub-procedures, in the main, tend to be reasonably reliable.

2. *VMS Sequences*. Many mathematical tasks have a certain quasi-periodic structure: one *sees* something, which leads to the retrieval and execution of some procedure; the execution of this procedure yields a modified visual input, which leads to the retrieval and execution of the next segment of procedures, and so on. One thus has a sequence of visual input, procedure execution, new visual input, procedure execution,...Such a pattern is more or less clearly recognizable in, say, a long division problem in arithmetic, or a "long" multiplication problem, or in the usual written algorithms for multiplying or dividing polynomials. [It is also familiar outside of mathematics, e.g., in driving somewhere if the route is complicated and is relatively unfamiliar. One hopes to recognize a landmark - oh, yes, at that school building one turns right - then to drive along until another key landmark triggers the next key decision and so on. One could not describe the entire route merely from memory, without the prompts given by key visual inputs along the route.] Sequences of this type have come to be called visually-moderated sequences ("VMS" sequences).
3. *Integrated Sequences*. Successful mathematics students, through practice, are able to develop many VMS sequences into holistic entities which they can contemplate *without* going through the entire sequence, step by step. (Similarly, our driver becomes able, with practice, to describe the entire route in detail - and hence can instruct others in exactly how to follow the route - without the need to physically re-trace the route in order to prompt his memory.)

Sequences which have reached this state are called *integrated sequences*. [For most present readers, long division is probably an example of an *integrated sequence*. One can "see the whole process in the mind's eye" - at least in general - without actually writing it down on paper. The specific numerals probably must be suppressed, but the general pattern can be visualized clearly.]

4. *Integrative Capabilities*. If VMS sequences can be transformed, by practice, into integrated sequences, there must be some mechanisms that make this transformation possible. In particular, there seem to be two easily-recognizable capabilities: a capability to identify similarities between two portions of procedures, and a "look-ahead" capability to recognize, before-hand, what is coming next. But since carrying out a procedure is itself an experience, quite probably *all* information-processing capabilities can be brought to bear upon *procedures* themselves.
5. *Record-Keeping Capabilities*. Clearly, there are record-keeping capabilities. Prominent among these seems to be a record of procedures actually used, as in Winograd's well-known Blocks World.
6. *Assimilation Paradigms*. This concept is drawn primarily from the work of Piaget (cf. Davis, 1971-2). It is historically earlier than the artificial intelligence concepts used in most of this discussion; because of this earlier date, the concept of *assimilation paradigm* is, unsurprisingly, more general, less specific than the A.I. concepts typically are. This lack of precision is useful, and is deliberately retained. Briefly: an "assimilation paradigm" is an idea in the student's mind that is retrieved in response to some information-processing "trigger," such as a cue in some visual input, and which then plays a central role in subsequent information processing. Thus, in the Piagetian "rotating jar" task, a young child (about age 3 years) may

draw a fluid in a jar *as if it were a frozen block of ice*, when in fact it is not; this is sometimes analyzed by assuming that the child retrieved an "assimilation paradigm" (AP) based on experience with solids, rather than fluids. Since the AP then influences subsequent processing, the child's drawings are essentially those of a solid, not of a liquid.

The question naturally arises, is an assimilation paradigm in fact the same thing as a *frame* (which is defined below)? Frames, drawn from A.I., are relatively precise and detailed; AP's are vague, general. This is a useful distinction to retain, at least for the present; it allows the making of less precise statements about AP's, or of more precise statements about *frames*.

7. *Frames, Scripts, or Schemata*. Although differences are distinguishable between the precise usages of Minsky, Schank, Rumelhart and Ortony, etc., no distinction will be made here, either in name ("frame," "script," or "schema") or in detail description. What is common to all three approaches is far more important than their differences.

Several considerations prompted the introduction of *frames*. For example, consider the behavior related to "understanding" a written passage, such as:

It was Paul's birthday. Jane and Alex went to buy presents. "Oh look," Jane said, "I'll get him a kite." "He already has one," Alex said. "He'll make you take it back."

Most U.S. adults, having read that passage, can answer questions such as:

- Q1: Why are Jane and Alex buying presents?
- A1: Because it's Paul's birthday.
- Q2: For whom are the presents intended?
- A2: Paul.
- Q3: Where did Jane and Alex go?
- A3: Presumably to a store that sells kites and other things children like.
- Q4: The next to the last word in the paragraph - "it" - to what does that refer?
- A4: To the kite that Jane was proposing to buy.

Having answered such questions, most humans are surprised to find that *not* one of those answers is given explicitly in the written passage. The answers have come from the reader's general knowledge of birthdays, presents, stores, etc. In Minsky's language, the reader "retrieved a birthday frame." *Frames* play an important role, because they embody a representation of a small part of the structure of the world. They are hypothesized to involve *variables*, with *default procedures* for assigning values to variables when input data fails to do so. [Asked to repeat the story a week or two later, most adults will unwittingly *add to it*, by inserting information from the "birthday frames" that was not included in the original story.]

There is a growing body of evidence suggesting the operation of some mechanism generally similar to "frames." The response to natural-language text is among the strongest. In addition to text studies like "birthday," consider also the well-known phenomenon of person A claiming that B told him fact X, when B denies having done so.

(For data specific to mathematics, cf. the study by Hinsley, et al., which we discuss below.)

8. *Frame Retrieval Mechanisms.* Clearly there must be some mechanism for retrieving an appropriate frame, if one can be found in memory. Some evidence exists that frame retrieval occurs early in many information processing operations, based on small early cues (cf. the Hinsley, Hayes, and Simon study, discussed further below).
9. *Frame Appropriateness Judgements.* After a frame is retrieved, it must be accepted or rejected. This corresponds, at least roughly, to a part of the "assimilation" vs. "accomodation" alternatives of Piaget.

10. *Assigning Values to "Variable Slots" in a Frame.* Clearly there must be some mechanism for determining an appropriate mapping of input data into the *variable* slots in a frame. This must be one of the most important information-processing functions.
11. *Meta-Language ("Planning Language"; "Planning Space").* Confronted by a problem, a student may need to retrieve one or more frames, and sequence them properly. Possession of appropriate *descriptors* (a "meta-language") for these frames can help. Consider the problem:

Find the distance from the origin  $O(0,0,0)$  to the plane through the points  $A(5,0,0)$ ,  $B(0,3,0)$  and  $C(0,0,2)$ .

This can be a hard problem, but is not, *if* one has both the necessary schemata for various parts of the problem, and appropriate descriptors to help get them selected and interrelated. If, for example, one knows that the dot product  $\vec{u} \cdot \vec{v}$  between any vector  $\vec{v}$ , and a unit vector  $\vec{u}$  in some special direction, gives us the *component* of  $\vec{v}$  in the direction of  $\vec{u}$ , one can find *any* vector from  $O$  to the plane (for example, the vector  $(5,0,0)$ ), and dot it into a unit normal vector. If one has any non-zero normal vector  $\vec{N}$ , one can get a *unit* normal vector by dividing  $\vec{N}$  by its length. If one has the equation of the plane in the form  $ax+by+cz=d$ , one easily gets a normal vector  $\vec{N}$ . And so on. Given *descriptors* such as these, the problem of selecting and sequencing the necessary schemata becomes entirely manageable.

B. *Some Theoretical Questions*

1. *To simulate or NOT to simulate?*

For many researchers, the value of AI analysis of human thought is that it can lead to actual simulation of human thought by computer programs. One advantage claimed, probably correctly, for creating simulation programs is that the process requires that every detail,

however small, be dealt with carefully, and that some of these "small details" may turn out to be crucial. There is also the further advantage that creating simulation programs provides a kind of "quality control" for cognitive theory, in the way that logic does for Euclidean synthetic geometry, that algebraic calculations do for Cartesian geometry, and that laboratory experiments do for physics and chemistry. "The moment of truth," Newell and Simon have remarked, "is a running program."

Creating simulation programs has NOT been part of the present research. We have sought, instead, to use AI concepts to create a *rhetoric* (or a collection of *metaphors*) appropriate for the analysis of mathematical thought by humans. In part this is a matter of aiming for a certain level of mathematical thought - namely, something of an "intermediate" level. We prefer NOT to be compelled to deal with such questions as:

- i) reading (interpreting) handwriting
- ii) comprehending ordinary English sentences
- iii) deciding which hand to use in picking up counters for "chip trading".

Instead, we *assume* the existence of sub-procedures capable of attending to many "details," and focus on the way that these sub-procedures can be combined in order to achieve relatively powerful results. In fact, this, too, can be accomplished by computer simulations, but the general tendency of simulation-writing is to be pushed toward dealing with minute details - and frequently details that are more characteristic of computers than they are of human thought.

- 2. *Are "Frames" Really Anything Distinctive?* The distinction between VMS-sequences, meta-analysis, and frame-retrieval is often characterized as the question: are key decisions made early in the consideration of a problem or is decision making distributed throughout the process of working on the problem? Our answer is: sometimes one, sometimes the

other depending upon the individual student and the specific problem.

[Cf., for example the task of integrating  $\int \sqrt{x^2+a^2} \, dx$ ].

Essentially this same conclusion was reached, independently, by Hinsley, Hayes, and Simon [1977]. Our conclusion came from observing thousands of episodes of problem-solving behavior. Theirs came from a particularly elegant sequence of five experiments. In one experiment, for example, it was found that, for some subjects, merely the noun phrase "A river steamer..." was sufficient to trigger the retrieval of an appropriate schema [indicated by the subject saying "It's going to be one of those river things with upstream, downstream, and still water. You are going to compare times upstream and downstream - or if the time is constant, it will be the distance."]. Hinsley, et al., study mainly older students than we do, and use easier problems, so that they more commonly observe mature schemata, but their results are very close to ours.

Are frames in fact really different from VMS sequences? We would state at least 5 points of distinction:

1. Frames allow major decisions to be made early in a process, as the study by Hinsley, et al., demonstrated quite dramatically. Merely the three words "A river steamer..." were sufficient to retrieve a complete frame for that whole category of algebra word problems. Everything that follows is then merely a matter of variable binding confirming correctness of frame and mapping choices, and executing explicit procedures that are referenced by the frame (or are part of the frame).

2. Because of the preceding property, frames make possible "top down" vs. "bottom up" processing. When a frame *is* retrieved early, it can guide us in looking for that input data which is most essential. Given 2 unknowns, we might seek two linear equations. Given a falling body problem, we can seek initial velocity, etc.

Indeed, having retrieved the "birthday" frame, we can ask: Whose birthday? How old? How many guests? What kind of party? and so on. All such seeking for input data to permit the binding of appropriate variables - all such "knowing where to look and what to look for" - is often called "top down" processing: the schema is itself guiding the data collection.

By contrast, one can have a more minute piece of data, not yet related to any specific "larger" schemata. It is a clue, not a *theory*. Followed by other clues, it can lead us through some tree, and hence determine the selection of an appropriate "larger" schema. This process is often called "bottom up" processing, and presumably represents part of the answer to questions about schema selection and retrieval.

The "top down" vs. "bottom up" possibilities require not only the existence of frames, but in fact the existence of a very large number of frames, of different "sizes," some small enough and specific enough to be the generalized cognitive representation of a single clue, and others large enough to represent a reality as large and internally structured as a "birthday" or "a falling-body problem."

3. *Frames* allow for "random access" variable binding, or other processing, whereas *procedures* are presumably more sequential. In "birthday," we can ask questions in almost any sequential order, and store the

answers via variable binding in the "birthday" frame. There are few, if any, restrictions on the sequential order of proceeding. We can ask first: Whose birthday? Or we can ask first: Will there be a party? Or we can ask first: Who are some of the guests? and so on.

4. Of course, of the greatest importance is the question of *default values for variables*. If I say "Leslie is a teacher," you can immediately retrieve one (or more) probably relevant frames, and can make immediate default evaluations. I can't predict what default value you have given to Leslie's sex (which, after all, *could* be either M or F), I can't predict whether your default values represent Leslie as a bearded guru wearing a white loin-cloth, or as a Hubert-Humphrey-like politician, or as a faculty member at Harvard Law School, or as a football coach, or as a kindergarten teacher. But you *are* able to make default evaluations, and to modify them as incoming data may require.
5. People probably display more confidence when their data-processing is based on frames than when it is based on VMS sequences or other purely sequential processes. The "authority" of a John Chancellor may come from the fact that his information processing is based on well-developed frames - we say "the news has meaning for him," "he knows what he's talking about." By contrast, a novice commentator would be "feeling his way," and would not give viewers the same impression of depth.
3. *Where Do Frames Come From?*

Our observations suggest a sequence generally like this: For novel problems something resembling VMS sequences will be learned (from books or teachers), which - given sufficient practice - will become schemata. Given yet more experience (and perhaps additional

explicit instruction), a meta-language will be created. This is the *general* pattern, though individual students may not complete the entire process, and simpler problems may not require it. [Note that the process is unendingly repetitive, and an elaborate solution initially synthesized by meta-language analysis may become a new VMS sequence, or a new schema.]

That is to say, as a VMS procedure is used, perhaps many times, as the *experience* of using the procedure is subjected to analysis by "look ahead" devices, pattern recognition mechanisms, and (presumably) the whole panoply of information analyzing devices that can be brought to bear on experiences in general, as the VMS procedure is acquiring the wholeness of an integrated sequence, as it is acquiring appropriate meta-language descriptors, it is in fact becoming a frame.

C. *Examples of Behaviors*

1. Erlwanger has found that the vast majority of VMS sequence errors made by the 4th, 5th, and 6th grade arithmetic students whom he has studied were made in super-procedures that were the subject of recent learning, and not in the sub-procedures which were called for by these super-procedures. [Erlwanger, 1975; Davis, 1977.]

A typical example:  $.3 + .4 = .07$

The sub-procedure that inputs 3 and 4, and outputs 7, performed correctly; the sub-procedure that counted decimal places and added them worked correctly; but, of course, this last sub-procedure ought not to have been activated for this problem.

2. Particularly interesting are the uni-directional "binary switches" [Davis, Jockusch, and McKnight, 1978], such as  $4 \times 4 = 8$ .

Errors of this type have the following attributes:

- i) Procedure B is used when Procedure A should have been used [in the example, addition was used when multiplication should have been];
- ii) The notations for Procedure A and for Procedure B are visually similar [ + vs.  $\times$  ];
- iii) The procedure *used* was learned earlier in the student's life than the procedure that was slighted - an *earlier* procedure erroneously displaced a "newer" one [addition is learned before multiplication in nearly all school curricula];
- iv) The switch (presumably because of (iii) ) is one-directional: The older replaces the newer, never conversely [we have taped no instances of  $4 + 4 = 16$ , or  $2 \times 3 = 8$  (but one often finds  $2^3 = 6$ )].

These uni-directional errors have an interesting theoretical explanation, in terms of the discriminations that are (or are not) made internally within procedures and super-procedures [cf. Davis et al., op. cit.]. Briefly, it is postulated that the discriminations which are made are no better than what is required at that time; hence, future situations arise (given new concepts, definitions, procedures, or notations) for which adequate discriminations have not been made a part of internal labelling processes.

3. Certain errors are extremely common and extremely persistent. These include:

- i) Distributive Law Errors, such as  
 $3(x^2 + y^2)$  rewritten as  $3x^2 + y^2$ ;
- ii) "Wishful thinking" errors, such as squaring

$$\left[ 2c + \sqrt{(x-a)^2 + y^2} \right]^2$$

to get

$$4c^2 + (x-a)^2 + y^2 \quad ;$$

iii) "Grouping" errors, such as:

(a) squaring

$$c \sqrt{(x-a)^2 + y^2}$$

to get

$$c^2 (x-a)^2 + y^2$$

(b) "simplifying"

$$(3x)^2 (-2y)^3$$

to get

$$9x^2 - 8y^3$$

[In example iii (a), the square root sign was playing two separate roles, an "exp  $\frac{1}{2}$ " role *and also* a "parenthesis" or "grouping" role; when squaring eliminated the need for the "exp  $\frac{1}{2}$ " role, *it did not eliminate the need for the grouping role.*

In example iii(b), it is true that  $(-2y)^3$  is equal to  $-8y^3$ , but once again the parentheses had been playing two roles, *only one of which is no longer needed.*]

iv) "Zero product principle" errors.

While it is true that

$$(x-2)(x+3) = 0$$

implies that either

$$x - 2 = 0$$

or else

$$x + 3 = 0,$$

it is NOT true that

$$(x-2)(x+3) = 8$$

implies that either

$$x - 2 = 8$$

or else

$$x + 3 = 8.$$

Because these particular errors are so extremely common, and so remarkably persistent, they deserve special study. *One suspects there must be powerful underlying mechanisms producing these errors.*

In the case of the "zero product principle" errors, an elegant analysis has been made by Marilyn Matz at M.I.T. [Matz, 1979]. Matz hypothesizes the existence of two levels of information-processing machinery:

- (i) a more superficial level that embodies the "rules" with which we are familiar, such as the long division algorithm, the column addition algorithm, the distributive law, the quadratic formula, etc.;
- (ii) a deeper level of "rules that make up rules." The zero product principle error is, in the Matz analysis, an error of a most familiar type - the insufficiently-quantified "overgeneralization" - but this error is made at the deeper level of "rules that create rules."

To begin with, recall the very common "overgeneralization" errors, e.g.:

Alex [aged 5 years]: "Daddy, I seed a butterfly!"

Alex has correctly perceived the regular rule for forming past tenses, but she has NOT yet learned the limitations on when this rule should be used. *This is one of the most common of all cognitive errors in children.*

Consider now this "deeper level of rules." One of them, Matz hypothesizes, is: "the numbers don't matter - it *could* be some other number: in fact, *any* number."

We have no doubt that such a rule must exist. As one example, an elementary school teacher in Berkeley, California, working with

$$\square + 3 = 3 + \square ,$$

thought about the problem, then excitedly exclaimed: "Oh, I see! That '3' doesn't have to be a '3'!"

Taken literally the statement is nonsense - a "3" that changed its value capriciously would have small usefulness, if any. But the meaning, clearly, is that

$$\square + 4 = 4 + \square$$

would be equally correct, and so would

$$\square + 5 = 5 + \square$$

or

$$\square + 10 = 10 + \square$$

or

$$\square + 1066 = 1066 + \square ,$$

or, in general,

$$\square + \triangle = \triangle + \square .$$

In short, she was using the Matz rule: *the numbers don't matter.*

Indeed, no one could learn arithmetic unless they did possess some such rule. Having learned that

$$3 + 3 = 2 \times 3$$

$$4 + 4 = 2 \times 4$$

$$5 + 5 = 2 \times 5$$

a successful student *must* infer that

$$6 + 6 = 2 \times 6$$

and so on. When we teach that

$$(x-2)(x+3) = 0$$

implies either

$$x - 2 = 0$$

or else

$$x + 3 = 0 ,$$

we hope the student will realize that "that '2' didn't have to be a '2' and that '3' didn't have to be a '3'." In short, we hope the student could also solve

$$(x - 5) (x + 4) = 0.$$

Putting together the *over-generalization error* and the deeper-level rule that "the actual numbers could, of course, be different; the numbers don't matter," Matz concludes that the extremely persistent *zero product principle* error owes its persistence, as dandelions do, to the fact that it is repeatedly re-created. Once eradicated, it is NOT gone for good - it will be created anew.

The deeper-level rule, observing that

$$(x - 2) (x + 3) = 0$$

implies that

$$x - 2 = 0$$

or else

$$x + 3 = 0,$$

concludes that, since the actual numbers don't matter, it must be true that

$$(x - 2) (x + 3) = 8$$

implies that

$$x - 2 = 8$$

or else

$$x + 3 = 8$$

The only "bug" in all of this is the failure to *restrict* the lower level rule to recognize the *exception* that is posed by the *zero* in the original example: *that this is one of the rare cases where the specific number DOES matter.* [Matz, 1979].

John Seely Brown has developed a similar analysis of certain subtraction errors [Brown, 1979].

4. An error common among beginning algebra students consists of combining like terms, as in

$$2x^3 + 5x^2 + x^3 + 2x^2 + 3x^3$$

to get

$$7x^2 + 5x^3 ,$$

treating the "invisible" coefficient *one* as if it were zero.

In fact, this is a predictable error, the consequence of using a procedure like this:

```

10  set n
20  identify next term  $x^n$ 
30  read off numeral to the left of  $x^n$ 
40  add into sum for coefficient of  $x^n$ 
50  loop back to line 20 as needed
60  reset n
70  go to line 20, unless done.
```

5. *Frame Errors.* One can easily find instances of frame errors of several types, but they can all be sorted into two broad categories: failure to retrieve a frame, vs. the retrieval of an inappropriate frame.

Central to questions of frame retrieval is the "top down" vs. "bottom up" distinction. When a frame has been retrieved and judged appropriate (at least *pro tempore*), information processing can be directed by the structure present in the frame. Having retrieved a frame for quadratic equations, we know to look for the coefficients, to see whether the discriminant is positive, negative, zero, etc., because the frame indicates the needed information. Having retrieved a frame for simultaneous linear equations, we know to ask whether the determinant of the coefficients is zero or not, because the frame asks for the data. Processes shaped by the demands of a frame in such a way are often called "top-down." The frame directs our processing, we "know where to look and what to look for."

By contrast, if we have some information processing task prior to frame retrieval, some bootstrapping operation must get us from "primitive" data to the stage of retrieving an appropriate frame. This is often called "bottom up" processing.

Presumably at the beginning of an information processing task - to the extent that a "beginning" can exist - processing must at first be "bottom up." One would not know, at first, which frame to retrieve (of course this can be modified by *mental set*, etc.). The transition to "top down" processing can occur very quickly, as Hinsley, et al., demonstrated dramatically. In one experiment, for example, it was found that, for some subjects, merely the noun phrase "A river steamer..." was sufficient to trigger the retrieval of an appropriate schema [indicated by the subject saying "It's going to be one of those river things with upstream, downstream, and still water. You are going to compare times upstream and downstream - or if the time is constant, it will be the distance."].

a. *Retrieval Errors - Failure to Retrieve*

Failure to retrieve any candidate frame whatsoever can occur if the individual *has no* appropriate frame stored in memory - surely one of the most common of all failures. It explains why most of us don't know most of the things we don't know.

But even if there *is* an appropriate frame stored in memory, one can still fail to retrieve it. B, a 9th grade student, was trying to solve a system of simultaneous equations in  $\cos\theta$  and  $\sin\theta$ .

After a protracted period of failing to solve the problem, B gave an exclamation, did solve the problem, and said "when I stopped thinking about trig and started thinking

about systems of equations" it became easy. Moral: one's memory is somehow partitioned. Searching the wrong sections will usually be unproductive.

- b. *Retrieval Errors - Retrieving the Wrong Frame.* But retrieving an inappropriate frame can be a particularly virulent form of bad news. P, a 10th grade student meeting differentiation for the first time, spent two weeks dealing with information processing tasks such as learning the formula

$$\frac{d}{dx} (u + v) = \frac{du}{dx} + \frac{dv}{dx} \quad (1)$$

and solving the problem of finding  $\frac{dy}{dx}$  if  $y = x^4 + 3x^2 + 8x$ , by retrieving frames dealing with variables, fractions, and the distributive law which were entirely unsuitable.

For (1), he assumed "d" and "x" were variables just like those he had dealt with before. Hence he assumed that (1) was nothing more than the distributive law. For the second problem, given

$$y = x^4 + 3x^2 + 8x$$

he divided by x to get

$$\frac{y}{x} = x^3 + 3x + 8,$$

then multiplied  $\frac{y}{x}$  by 1 in the form of  $\frac{d}{d}$ , to get the (wrong) result

$$\frac{dy}{dx} = x^3 + 3x + 8.$$

In the sequence

$$\begin{aligned} \frac{dy}{dx} &= \frac{d}{dx} (x^3 + 5x + 2) & (1) \\ &= \frac{d}{dx} (x^3) + \frac{d}{dx} (5x) + \frac{d}{dx} (2) & (11) \end{aligned}$$

$$= 3x^2 + 5, \quad (iii)$$

whence  $y' = 3x^2 + 5, \quad (iv)$

P objected that the  $\frac{d}{dx}$  had gotten lost, and should appear in line (iv).

When circumstances finally compelled him to recognize that his whole approach was quite wrong, he was, unsurprisingly, devastated. The retrieval of an inappropriate frame - and its use for two weeks - had left him far behind his classmates. There must have been a growing accumulation of doubt-casting pieces of information, but for two weeks he had continued, presumably by heroic efforts, to try to fit everything into his purely algebraic frames.

One can get an empathetic comprehension of P's plight by imagining a student confronted by the Cayley-Hamilton theorem, which states that the matrix A satisfies its own characteristic equation,

$$\det (A - \lambda I) = 0. \quad (v)$$

If one replaces  $\lambda$  by A in equation (v), it seems that the Cayley-Hamilton theorem is a complete triviality:  $AI=A$ , and  $A - A$  is the zero matrix, and the determinant of the zero matrix is surely 0. What's all the fuss about?

Nothing blocks the search for an appropriate frame more effectively than the retrieval of a frame that (perhaps incorrectly) passes the adequacy test. Those who are satisfied do not seek.

- c. *Retrieval Errors - Semantic Templates.* R, an adult studying college algebra at a community college, was confronted by the

need to compute

$$(\sqrt{x+a} - \sqrt{x-a})^2$$

by carrying out the squaring. When the instructor referred to  $(A+B)^2$ , R considered this an irrelevant digression, and did not see how to make use of it. This was a typical response by R. Apparently R has not recognized that one has certain key patterns in one's mind - call them *semantic templates* - and that mathematics sometimes requires an effort to match up some notation on paper with an appropriate semantic template. R feels that something explicitly present in

$$(\sqrt{x+a} - \sqrt{x-a})^2$$

should determine the next step. In effect, she expects all of mathematics to resemble the long division algorithm, to be a rote VMS sequence.

A slightly similar, though far less extreme, situation is readily observable among tenth grade students trying to find the equation of an ellipse from the "sum of the distances from two fixed points" definition; one easily gets, by a step-by-step process, to the result

$$a^2x^2 + c^4 = c^2x^2 + a^2c^2 + c^2y^2. \quad (vi)$$

Nothing in this line clearly determines what should be done next. To get further, one must shift gears and begin to think of the desired result. If one knows that the goal is an equation of the form

$$\frac{x^2}{R^2} + \frac{y^2}{W^2} = 1, \quad (vii)$$

for some R and some W, then getting from equation (vi) to the final result is easy - but what is required is a change from step-to-step processing, switching instead to a

strategy of keeping the goal clearly in mind.

- d. *When You Do NOT Want Frames.* To some extent one ordinarily considers a frame-like organization of knowledge to be more mature, and more secure, than sequential organization. Of course this view overlooks the cyclic reality of small pieces being "chunked" together [Miller], thus getting a new collection of primitive elements, soon to be in need of further chunking.

It also ignores another reality: sometimes rigid sequential processing is to be preferred as more appropriate than frame-like or "gestalt" processing.

E, a 10th grade geometry student of high ability, for whom English was not a native language, had studied the focus-directrix definition of the parabola, and the two-focus definition of the ellipse and hyperbola. On an examination the class faced a novel task: finding the equation of the paths of a point that moves so that its distance from a fixed point is always half its distance from a fixed line. The rest of the class proceeded as expected: writing the equation of distances, using the distance formula, squaring, and simplifying. At this point the best students were surprised; they had expected a parabola (from the "point-line" form), but recognized that they had obtained the equation of an ellipse. E proceeded entirely differently. She identified specific points on the path, recognized that the path was apparently an ellipse, wrote (from memory)

$$\frac{x^2}{c^2} + \frac{y^2}{b^2} = 1,$$

and found  $c$  and  $b$  by considering recognizable points on the ellipse.

What had E done? In effect, instead of using a step-by-step procedure, she retrieved a frame, made appropriate replacements for frame variables, and proclaimed the result.

This was not what had been intended, and was in fact not a proof, yet it was not entirely without merit: guessing the form of an answer, and subsequently determining the details (such as specific values for variables) may well be the most common general method in all of applied mathematics.

#### 6. *Language-Related Errors.*

Certain errors are seemingly related to ambiguities in, or poor use of, natural language.

Jennifer, a third grade girl, claimed that the written expression

804

and the written expression

84

were the same, because "adding zero makes no difference; zero means nothing." She similarly argued for no difference between

8

and

80

and

800 ,

when all three were presented to her simultaneously in *written* form.

Similarly, a common error among beginning algebra students is to pronounce

$$2x^2 - x^2$$

as "two x squared, take away x squared." Students who read it this way nearly always write the result as

$$2x^2 - x^2 = 2 .$$

7. *Epistemological Errors.* Davis, Jockusch, and McKnight list a collection of "epistemological" errors. We cite one type. Students notoriously have concerns about the propriety of defining 0! as 1, or of defining  $A^{\frac{1}{2}}$  as  $\sqrt{A}$ , or of defining  $A^{-1}$  as  $\frac{1}{A}$ . This difficulty comes mainly from a failure to recognize that *they are building up a system of mathematics*. At any given point the system is only partially completed. There are "frontiers" where further definitions have not yet been made (although appropriateness criteria for future constructions may be implied by what has already been built).

This is a special problem for younger students, presumably because they have not yet lived long enough to realize that their city, their state, their nation, their laws, their traditions, their literature, their culture - and, indeed, themselves - are still being built. None is yet finished.

8. *The Simon-Newell Phenomenon.*

In observing students, we have been forcefully struck by the fact that most of the best problem solvers solve problems - even relatively novel ones - with great speed and seemingly little analysis. They "just know how to solve them." It is somewhat akin to the way the reader might find the sum

$$79 + 2 = ?$$

very quickly and with no conscious effort at analysis. By contrast, *after* they have solved a novel problem, strong problem

solvers are often lost in thought for surprisingly long periods of time. It is as if the truly *creative* process occurs *after* you have solved a novel problem, and are remorselessly pursuing questions about what God hath wrought, making the associations, reconciling the seeming contradictions, identifying the common patterns that will fit this new knowledge securely (and fruitfully) into a much larger body of knowledge. [Incidentally, we believe we are finding references to this phenomenon in Polya, Kilpatrick, and Newell, and it is clearly implied by recent work of Simon and several of his students. Consequently, we have come to call it the "Simon-Newell Phenomenon."]

9. *Meta-Language Analysis, or "Planning."*

- (a) We have referred earlier to the novel problem of finding the distance from (say) the point  $P(a,b,c)$  to the plane that passes through the three points  $R(r,0,0)$ ,  $S(0,s,0)$ , and  $T(0,0,t)$ . This requires some form of planning language.
- (b) One can look explicitly for schemata or meta-language analysis in student performance. For example, Davis et al. [1978] reports interview protocols that indicate that the most able ninth graders observed (who had studied some calculus intermixed with their ninth-grade algebra) were able to deal with a complicated VMS sequence not merely as something to do (a "verb"), but also as a thing to discuss and to analyze (a "noun"); this process is aided somewhat if an appropriate meta language is provided (though the most able students can create a meta language without help). For example, consider the problem: A ball is thrown straight up so that it reaches a height of 40 feet. How fast is it thrown? The weakest students cannot solve this problem. Somewhat

stronger students can solve it, by carrying out an appropriate VMS sequence, following procedures demonstrated in class, but cannot *discuss* the problem abstractly *without actually carrying out each step*. The (few) strongest students can discuss and analyze the solution correctly, in detail, without carrying out each step. For example, somewhat paraphrased, here is an actual protocol from a ninth-grade student of high ability.

"From  $h''(t) = -32 \text{ ft/sec}^2$ , you can get  $h'(t, V_0)$  with an unknown constant  $V_0$ . But for  $h$  to be a maximum,  $h'$  must be zero. The equation  $h'(t_{\max}, V_0) = 0$  gives  $t_{\max}$  in terms of  $V_0$ . Integrating  $h'(t)$  gives  $h(t, V_0)$ . Now, for  $t = t_{\max}$ ,  $h(t_{\max}, V_0) = 40$ , and since  $t_{\max}$  can be expressed in terms of  $V_0$ , we can solve for  $V_0$ ." [This is not verbatim, but does follow the answer of the student, talking accurately, in detail, about how he would solve the problem - *without executing* the steps of the actual VMS sequence, and hence without benefit of the sequence of visual cues as the written solution develops. The student's actual language made use of phrases such as "plug in this expression for  $t_{\max}$ ," and so on.]

10. We have explored procedures, especially the relation of super-procedure to sub-procedures, in looking at the learning of some of the more algorithmic aspects of school algebra. One technique in those explorations might be called the "before and after snapshot." Brief (about five minutes) quasi-clinical interviewing was used to generate talk-aloud protocols for students solving two or three routine problems and to gather student clarifications of ambiguous points in the protocol. "Before" and "after" pictures were obtained by conducting two interviews on the same type of problem, the first within the first day or so of instruction relevant to the problem type and the second towards

the end of immediately relevant instruction. One such "before and after snapshot" is discussed here, but we have collected many such protocols, of which this one is fairly typical.

The subject was a seventh-grade boy, whom we shall call "P", who had been in an introductory algebra course for about two and a half months. The course had covered signed numbers, some beginning analytic geometry, logic and formal derivations of algebraic expressions and it was just then moving into instruction on simplifying algebraic expressions prior to going into the arithmetic of polynomials. Thus, tasks like

"Simplify  $4x^2 + 2x - [x^2 + 3x]$ "

"Simplify  $5x^2 + 3x - [x^2 + 4x]$ "

represented new, challenging problems although some preparation had been made during the course of instruction on algebraic derivations (which involved use of formal axioms like the distributive law, commutative law of addition, etc.)

Let us first look at P's "after" picture which came on the eighth day of instruction on the topic just before moving to the next topic. Given the task, "Simplify  $5x^2 + 3x - [x^2 + 4x]$ ," his protocol may be summarized as follows:

- (1) He first announced that he was going "to make the minus into plus a negative" and wrote

$$5x^2 + 3x + -[x^2 + 4x].$$

[Notice, he unhesitatingly calls up and executes this simple syntactic transformation procedure. He possessed a verbal label for what he did, had no doubts that it was a legal syntactical move, and knew that it would be helpful to him overall, suggesting that he has a clear super-procedure embedding this step.]

- (2) Next he announces that he is going to handle the "parenthesis" and writes

$$5x^2 + 3x + -x^2 + -4x.$$

[Again, a clearly legal syntactic transformation based on the distributive law which P executes without question or hesitation, indicating the functioning of an organized super-procedure. Instruction had suggested this transformation as one of several that would be legal and effective in what, in meta-language, would be called "removing parentheses." Stress was placed on a meta-language discussion of the super-procedure rather than extensive drill on executing one form of it. Notice P calls the brackets here "parentheses" even though the instructor and the interviewer had called them "brackets." The meta-language discussion has "taken," at least in part.]

- (3) P next uses the commutative law of addition twice to bring his "like" terms together, writing

$$5x^2 + -x^2 + 3x + -4x$$

[Again, he does this unhesitatingly and without pause between the two executions, indicating a clear super-procedure. He mentions that he can now do this since it's "addition," which suggests he has some awareness of why the first two steps were necessary. The effect of formal derivations on P's procedure is interesting. Further work typically curtails a step like this one, leaving it implicit in the step which follows, "combining like terms." Thus, this step indicates that the structure of P's super-procedure has been shaped by his experience in formal algebraic derivations in which he has made moves like this numerous times. This is strange in comparison to (1) the fact that, asked to name the justifying law here, he comes up finally with the associative law rather than the commutative and (2) the confusions in his "before" protocol which will be discussed below.]

- (4) Finally, he twice adds "like" terms, writing the final result,

$$4x^2 + -1x$$

[Again, clear super-procedure and correct sub-procedures. Interestingly, the final "-1x" in his answer reveals the carefulness and precision of his syntactical transformations. Such carefulness is also curtailed with further experience and confidence. Again, such clarity about syntactical matters will be seen to be in marked contrast to his "before" protocol.]

This "after" protocol seems to be obvious and to be a simple exercise in rhetoric application. The usefulness of the approach becomes clear in contrasting the "before" protocol to this precise, efficient, correct target performance in the "after." The "before" protocol was taken at the beginning of the third (45 minutes) period of instruction on the topic. P had worked simpler exercises and had seen worked exercises something like his assigned task, "Simplify  $4x^2 + 2x - [x^2 + 3x]$ ." His "before" protocol may be summarized as follows:

- (1) P announces that he wants to do something with the "like terms," identifying  $4x^2$  and  $x^2$  as such, but he is concerned that "it's minus that [ $x^2$ ] so...?" Later he expresses confusion about the fact that the 4 and the  $x^2$  in the  $4x^2$  are combined by multiplication while the 3x and the  $x^2$  in  $x^2 + 3x$  are combined by addition. [Notice, P is at first unable to call any sub-procedure; his "so..." trails off into a question. He has no super-procedure ready at hand. At best, he seems to have some heuristic guidance in the form of knowing he needs to "combine like terms." This leads him to reflect on the expression and his reflecting is a strange mixture of syntactic awareness and confusion. He realizes the minus should tell him something but he's not sure what. He contrasts the connectives in  $4x^2$  and  $x^2 + 3x$ , which is astute in the sense that most similar students would not have that clear a grasp of the syntax of  $4x^2$  or be led to compare the two sub-expressions. Yet, the contrast is almost esoteric in its irrelevance to the needed procedure and shows a confusingly wide-ranging, undirected reflection.]
- (2) After noting "you can't add," he copies down  $4x^2 + 2x$  without combining unlike terms as he briefly considered, and then goes on to complete the expression,

$$4x^2 + 2x - 3x + x^2$$

[P's explicit heuristic of combining "like terms" (apparently with a further implicit heuristic of bringing "like terms" in proximity as we see correctly done in a full-blown form in step 3 of the "after") and his syntactical confusions lead him to take the plunge with an incorrect, *ad hoc* procedure. He has the syntax of  $4x^2 + 2x$  correct and even the commutation of  $x^2 + 3x$  would be correct *inside parentheses*. He seems to have no awareness of the syntactical function of the parentheses nor any legitimate procedure for removing them. Thus the *ad hoc* move that is guided only by rough heuristics, not an organized super-procedure, and which shows lack of knowledge of a key syntactical point or else a lack of use of such knowledge in a validity check of the simple *ad hoc* procedure he has concocted (to relieve the pressure on him?).]

- (3) P next said that he has to "take" the 3x from both the 2x and the  $4x^2$ . He then wrote a -1 below the 2x and indicated that  $2x - 3x$  gave -1. When questioned further, he stated that "you minus the x [in 3x] from 2x." [Having now followed his rough heuristics to the point where they can now dictate directly a step, P proceeds to follow up his syntactical blunder in the preceding step with a further syntactical blunder. This is a clear case of some sort of *regression* under strain because in previous days P had easily done related simpler problems correctly using the same procedure which would be the appropriate sub-procedure here. Thus, encountering sub-problems which he

should be able to do in a context that is already demanding much of his cognitive resources causes malfunctions in his syntactical understandings and in his sub-procedures.]

- (4) P now moves confidently to take  $3x$  from  $4x^2$ , writing  $1x$  below the  $4x^2$ . Then he suddenly expresses some doubt that this is correct and, after a pause, decides it is not, because " $x^2 - x$ " doesn't "equal  $x$ ." He then goes back to consider whether he needs to take the  $3x$  from both terms and eventually decides he does not. [The confidence with which he begins this step seems to show that P has at least a rudimentary super-procedure to cover a few steps. The lack of hesitation seems a clue to such super-procedures. Suddenly his rough heuristic about "like terms" seems to enter in a different way, acting as a validity check which leads him to reject what he has just done. It is not clear why he is able to go on to realize that he was making a previous error in even taking  $3x$  from  $4x^2$ . A reasonable hypothesis, based on previous evidence of the good level of knowledge P has about algebraic syntax, is that it is this syntactical knowledge that surfaces here. We have often seen that one error will start children on a chain of errors like some short-term, localized algebraic dementia. P's earlier wrong transformation seems to have done this, but encounter with some syntactical fact of which he is sure brings him out of it and leads him to correct another of his errors.]

- (5) P then rewrites  $4x^2 + 2x - 3x + x^2$  to give

$$4x^2 - x^2 + 2x - 3x$$

[Somehow, "like terms" has put him back on the right track and he can access and use the same sub-procedure that he uses in step 3 of the "after" and, if the previous steps had been correct, it would be appropriate here. Unlike step 3 of "after," this sub-procedure is not executed as a part of a super-procedure but rather as a part of something like a "visually moderated sequence" (Davis, 1979A & 1979B) in which visual and other outputs from a previous step lead to selection and execution of a sub-procedure even without an overall plan.]

- (6) Finally, P moves quickly to write 4 as the output of  $4x^2 - x^2$  and  $-1$  as the output of  $2x - 3x$ , giving him a final answer of  $4 + -1$ , which does not trouble him at all. [This, of course, is the application twice of the same inappropriate procedure seen in step 3 of this protocol. It is mysterious why P can snap out of pressure- or error-induced regression enough to correct two errors but not go on to correct a third. It certainly goes to show the persistence of some errors. One possible difference is that while P has performed a sort of validity check both in step 3 and in step 5, that of step 5, " $x^2 - x$ " is correct and shows his error while that of 3, "you must minus the  $x$  from the  $2x$ " is a syntax-ignoring, misleading verbalization. Such a verbalization may have short-circuited the validity check process and left the error-producing sub-procedure intact.]

While P's final answer is, of course, less than satisfactory, and while he showed several shortcomings in this "before" performance, his actual

errors are few and easily corrected. It simply took pointing out the error in steps 3 and 6 to make P realize that it was wrong and, even to him, silly [although other students have proved more resistant to such knowledge], and a brief bit of instruction about the syntax of parentheses and about their removal. At least one practical benefit of the super/sub-procedure analysis is to pinpoint errors sufficiently precisely to greatly aid in remediation.

Comparison of P's "before" and "after" protocols suggests several conclusions (which are supported by other related protocols) including the following:

(1) The key task in tasks like that of P is really that of *super-procedure formation*. Probably the strongest contrast between before and after is that of the confidence and lack of hesitation in the after which is justified by the correct sequence of sub-procedure accessing.

(2) Execution of already formed procedures as sub-procedures in such a context (i.e., super-procedure formation) often results in a *decay* in performance of the already-formed procedures. This is similar to what Bruner and others have found in motor skills learning. It may make a case for over-learning of procedures that will be sub-procedures in a variety of more complex procedures.

(3) The first two phenomena combine in a change which might be called *mutual clarification*, i.e., when a correct super-procedure is stably formed, so will be correct sub-procedures. While it was not wholly the case with P, in some cases we have seen group instruction which continued to focus on forming the correct super-procedure result in seemingly spontaneous correction of sub-procedures.

(4) Formation of super-procedures seems clearly to run from trial-and-error crystallization around a few *heuristic* guides to a stable, *stereotypic* structure. This might be called an exodigmatic (paradigm from "within") formation approach as contrasted to an hypodigmatic (paradigm from "above") approach (i.e., repeated imitation of an overall model).

(5) The *syntax* of algebraic expressions may be a key or milestone kind of knowledge in algebraic learning. The degree of syntactic security seems to be a crucial element in a student's predisposition to regression under strain.

(6) Formal *derivations* (such as these students had been exposed to) may help in creating secure syntactic awareness and transformations even though it seems often to fail at the task of making the student able to provide explicit justifications.<sup>5</sup>

---

5

One is struck by the similarity of this interviewer/observer description of P's performance, and Small's description of rats running mazes:

In appreciating the results of this series of experiments, about the same facts come into view, only more distinctly, as in the case of the wild gray rats; the initial indefiniteness of movement and the fortuitousness of success; the just observable profit from the first experiences; the gradually increasing certainty of knowledge indicated by increase of speed and definiteness; and the recognition of critical points indicated by hesitation and indecision; the lack of imitation and the improbability of following by scent; the outbreak of the instincts of play and curiosity after the edge of the appetite is dulled. In addition are to be noted the further observations upon the contrast between the slow and cautious entrance into, and the rapid exit from the blind alleys, after the first few trials; the appearance of disgust on reaching the end of a blind alley; the clear indication of centrally excited sensations (images) of some kind; memory (as I have used the term); the persistence of certain errors, and the almost automatic character of the movements in the later experiments. [Small, 1901, p. 218; quoted in Olton, 1979, p. 58] \*

Olton's analysis of the mental processes of rats is often based on a conceptualization of "procedures" quite similar to the approach we have used in our research. Cf., for example, his use of *procedures* to explain the "response strategy" vs. "place strategy", where in our notation, he considers the possibility that the rat learns conditional rules:

If start ← A, turn right.

If start ← B, turn left.

## 11. "Heuristic" Frames

As a final example, we are coming to believe that one can sometimes distinguish "performance" frames from "heuristic" frames. Given the problem: prove that two lines  $l_1$  and  $l_2$  are parallel if and only if their slopes,  $m_1$  and  $m_2$ , are equal, the first step in several protocols was to prove that, if  $m_1 = m_2$ , then  $l_1 \parallel l_2$ . But for the second part of the proof one found two quite different approaches. One 10th-grade student assumed  $l_1 \parallel l_2$ , constructed a "horizontal" transversal, identified congruent angles, constructed congruent triangles, and computed the slopes of each line from these congruent triangles, getting  $m_1 = m_2$ .

By contrast, an adult mathematician used the device of choosing a point  $P$  on  $l_2$ , constructing through point  $P$  a line  $l_3$  with slope  $m_3 = m_1$ , and citing the parallel postulate to establish that  $l_3 = l_2$ , whence  $m_2 = m_3 = m_1$ . He described this as the strategy: "introduce a new entity  $E_3$  having the properties that you hope some entity  $E_2$  has, and then proving that  $E_3$  is actually  $E_2$ . He used this strategy in a wide variety of problems, where  $E_3$  might be a line, or problems where  $E_3$  might be a triangle, etc.

We are inclined to classify the " $E_3$ " strategy as "the use of a

## 5 Cont.

This kind of analysis, not unusual in 1979, was essentially unknown in 1946-7, when earlier studies of this type were conducted.

The "look ahead" operators which we hypothesize were familiar as early as 1932, although they were given a kinesthetic meaning irrelevant to their essential information-processing role. (Of course, in 1932 one did not speak of "information-processing roles," but it was quite common to speak of "kinesthetic impulses.") Cf. this example:

the general assumption underlying the function of kinesthesia in learning is that stimuli arise in some sort of temporal sequence from the activity of the muscles as the animal traverses the maze pattern. The further assumption which must be made is that during the process of learning this temporal sequence becomes fixed...The "movement to come" is released at the proper time by the afferent "kinesthetic" impulses aroused by the movement which has just been made. [Dorcus and Gray, 1932, p. 447; quoted in Olton, 1979, p. 585] \*

\*Copyright 1979 by the American Psychological Association. Reprinted by permission.

*heuristic* frame," contrasted with the 10th grader's strategy, which we would describe as "the use of a *performance* frame." Both seem frame-like, but the 10th grader's solution was more complicated, and made use of very specific, detailed work that had been done during the year. We were struck by what seems to be *the greater simplicity* of the adult's solution - not what we had initially expected. But perhaps it is what one ought to expect.

#### VI. *Corollaries and Implications*

This section deals with work in progress. Clearly one of the major goals of this project has been to create a background for further studies of students learning mathematics - a background that would give investigators a better notion of what to look for, and a better framework for interpreting what they observe. We report briefly on where things stand at present.

##### A. *Focused Inquiries*

A careful program of observation and description will inevitably move into occasional observations of a more sharply-focused nature. This process is now underway in our own research, and we are undertaking seven small studies of a more sharply-focused type, none of which could be completed within the present project, though a *first-pass treatment of one of them has been completed, and is reported in full detail in Davis, McKnight, Parker, and Elrick, 1979.*

##### 1. *Frames vs. Processing Strings of Symbols*

In the study of *reading*, it is common to distinguish those word recognition processes which occur slowly and in recognizable stages - such as "phonics" or "sounding the word out" - vs. those word recognition processes which occur almost instantly and tend to defy any analysis into component sub-parts (known as "look-say" or "sight word" recognition).

Similar distinctions seem to exist in the processing of mathematical notations. This led to (among other things) an experimental study of one class of fifth grade students of diverse SES and racial/cultural backgrounds.

These students had not yet studied negative numbers. A test on negative numbers, devised for seventh-grade students by K. Tatsuoka and her colleagues, was programed on the PLATO CAI system and administered to this single class of fifth graders. For details, see Davis, McKnight, Parker, and Elrick, 1979.

The essential point is this: consider the problem

$$4 - 5 = \underline{\hspace{2cm}}.$$

A student who knew about negative numbers would presumably respond correctly, with -2 or  $-2$ , read as "minus two" or "negative two," depending upon which books he had studied.

But consider a student who had *not* studied negative numbers. Our analysis of primary grade arithmetic frames has convinced us that most students develop a *frame for symmetric subtraction*, that might be represented like this:

*Subtraction*

This frame requires two input numbers.  
GET THEM!

one input number:                      (A)  
other input number:                      (B)

Return difference of A and B.

Although of course young children do not *think* of the process in such terms, in fact this frame returns the *absolute value* of the difference,

$$|A-B|.$$

Recalling the Hinsley, Hayes, and Simon result that frames may be retrieved quickly, from very scanty cues, one would then hypothesize that many students, upon seeing

$$3 - 5 = \underline{\hspace{2cm}},$$

would pick up the subtraction sign cue, retrieve the frame for symmetric subtraction, and consequently answer "2."

On the other hand, a student who attempted to process the symbol string 3-5 carefully, one symbol at a time, by referring to the *meaning* of subtraction, would presumably conclude that the problem is impossible.

In the PLATO experiment in question, the computer itself collected the data (the only time we have used this method of data collection; ordinarily human observers would watch the children at work, and interview them). The student answers were used to determine which method each student used on each problem. *The vast majority of student answers indicated use of minimal CUES to retrieve primary grade arithmetic frames, such as the frame for symmetric subtraction.*

A second example may make the distinction somewhat clearer: consider the problem

$$(+3) (-5) = \underline{\hspace{2cm}}.$$

A student who knew the arithmetic of signed numbers would presumably answer, correctly, either -15 or -15 ("minus fifteen" or "negative fifteen"), depending upon which book he had studied.

A student who used minimum-cue frame-retrieval processing would have two choices:

1.   i)   ignore the parentheses (which are entirely unfamiliar)
- ii) observe the "minus" sign, retrieve the primary grade frame for symmetric subtraction, and answer "2";

or else

2.   i)   ignore the parentheses (as in (1))
- ii) observe the "plus" sign, retrieve the primary grade frame for addition, and answer "8".

A student who attempted to read the string of symbols, and to make use of the *meaning* of subtraction, might conclude that "taking 5 things away from 3 things was impossible," and report that the problem could not be solved.

The phenomenon of students ignoring the parentheses is a common one, and a general one. (Notice that fourth or fifth grade students who are prolific readers will often claim that they have never seen semi-colons [presented visually as ";"], although in the course of their reading they have encountered thousands.) Herbert Simon<sup>6</sup> has pointed out that this "invisibility" is

---

<sup>6</sup> Personal communication to the authors.

explainable in terms of *production systems*: conditional instructions that would be triggered by parentheses have not yet been synthesized; the phenomenon can also be explained by assumptions of "top down" processing under control of primary-grade arithmetic frames, which never call for input data involving parentheses.

## 2. *John's Remarkable Facility*

One mystifying case has acquired a renewed interest because we have recently seen a new possibility of explaining it.

John, a student observed for three consecutive years, beginning in grade 7, typically exhibited great facility in finding correct interpretations for novel mathematical situations. This struck his teachers, intuitively, as almost a "linguistic" facility - he seemed to be very quick with the language of mathematics. No adequate explanation of this facility was ever found, nor was it even possible to find an adequately sharp definition of his skill.

He was NOT one of the most powerful mathematics students; on the contrary, he used his facility to progress with quite modest expenditures of effort. In Section 3, immediately following, we present a formal theory that *may* shed some light on John's facility, but the theory was developed too late to be applied to John's case. (In fact, the theory grew out of observations of Anne and several other *very unsuccessful* students.)

## 3. *Imitation vs. the Use of Semantic Knowledge*

The theory presented here has three points of origin.

- 1) Anne, and other students observed in the course of this study, used a strikingly "mindless" or "thoughtless" form of *imitation* to learn mathematics. The long-term effect was severely limiting - none of these students progressed satisfactorily.

- ii) How can statements in natural language - e.g., English - be used to guide mathematical operations? We have seen (e.g., in the case of Jennifer) that natural language statements can be a source of serious difficulty, as when "zero means nothing" and "adding zero doesn't change anything" were used to justify the claim that

80

and

800

were the same number. But obviously a sizeable amount of knowledge, much of it reliable, does seem to be represented as natural language statements.

The problem is really more general. Algorithmic or procedural representation is central to the conceptualization developed here, and the conceptualization seems generally able to deal with knowledge represented in this form. Knowledge represented in natural language statements is something else again - for the most part the present conceptualization doesn't even *attempt* to deal with knowledge represented in this form. But there are still other forms. When children use "manipulatable physical materials" - such as Cuisenaire rods, Dienes MAB blocks, chip trading, etc. - to learn about mathematics, one has a kinesthetic/visual/tactile experience that may be represented in memory in yet other ways. And it is often believed that there is some form of *geometric* representation for some knowledge.

When the human mind seeks to carry out some essentially

algorithmic task, how - if at all - does it make use of knowledge stored in memory but represented in these *other* forms?

- iii) The third - and most fruitful - point of origin for this theory is the work by Simon and his students in simulating learning to solve equations by studying the illustrative examples given in Chapter 8 of Stein and Crabill (1972).<sup>\*</sup> Because this has proved the most fruitful approach, we begin our discussion from this perspective.

Example 1.

$$3x + 4 = 0$$

Stein and Crabill present a solution as follows:

| Step                          | Reason                                      |
|-------------------------------|---------------------------------------------|
| $3x + 4 = 0$                  | Write the given equation                    |
| $3x + 4 + (-4) = 0 + (-4)$    | To undo addition by 4, we subtract 4        |
| $3x = -4$                     |                                             |
| $\frac{3x}{3} = \frac{-4}{3}$ | To undo multiplication by 3, we divide by 3 |
| $x = \frac{-4}{3}$            | Cancel the 3's                              |

We shall follow the illustrative example only as far as the step

$$3x = -4$$

As we follow this, keep in mind:

- a) A computer program that is "watching" what the textbook does, and trying to "program itself" by creating rules that will enable it to solve other, similar problems. [Of course, the computer program does NOT know which other problems deserve to be called "similar."]
- b) Anne, and other similar students, who try to learn algebra by "thoughtless imitation," and *do* succeed in imitating, but do *not*

<sup>\*</sup>From ELEMENTARY ALGEBRA: A GUIDED INQUIRY by Sherman R. Stein and Calvin D. Crabill. Copyright 1972 by Houghton Mifflin Company. Reprinted with permission of the publisher.

succeed in really learning algebra.

The textbook, faced with the input

$$3x + 4 = 0$$

responds by doing this:

$$3x + 4 + (-4) = 0 + (-4)$$

The computer (or Anne) now has the job of figuring out what on earth the textbook just did. The present reader, who knows *what* the textbook did and *why* it did it, may have trouble recognizing the complex problem which Anne and the computer now face.

There are, after all, *many different things that the textbook could have done, ALL OF WHICH WOULD LOOK EXACTLY THE SAME IN THE PRESENT EXAMPLE.*

At perhaps the least structured level, perhaps the textbook's strategy is to add randomly-chosen numbers at randomly-chosen points within the symbol string, and just *happened* to pick these points in the symbol string for insertions

$$3x + 4 \quad \uparrow \quad = 0 \quad \uparrow$$

and just *happened* to pick the number -4 for the first insertion, and the number -4 for the second insertion.

If that is a correct analysis, then one could imitate the textbook by dealing with the problem

$$5x + 7 = 0$$

by selecting insertion points, perhaps these three

$$5 \quad \uparrow \quad x + 7 \quad \uparrow \quad \uparrow \quad = 0$$

and selecting the numbers -1, 19, and 1066 for insertion:

$$5 + (-1)x + 7 + 19 + 1066 = 0$$

Again, the *reader* knows that this is NOT a correct "imitation" of the textbook's method - *but how is the computer to know?*

There are an unlimited number of possible "imitations." Despite this

awe-inspiring complexity, we can describe a few of the *possible* methods for "imitation" - with or without using various other forms of knowledge.

Method  $\alpha$  This method recognizes the two numbers 3 and 4 as inputs, *but does not distinguish between them.* [It selects inputs as if from an urn, with replacement.] It does not distinguish the two sides of the equation.

Hence, an "imitation" based on Method  $\alpha$  could look like this:

$$3x + 4 = 0$$

$$3x + 4 + (-3) + (-4) = 0$$

Method  $\alpha$  selected a number from the "input urn," getting 3; returned the number to the urn, made a second selection, getting 4; chose a side of the equation (getting "left"); replaced "left" in the urn, drew again, and again drew "left." From 3 it computed  $(-3)$ , and from 4 it computed  $(-4)$ ; the final result, then, would be

$$3x + 4 + (-3) + (-4) = 0$$

Of course, this same identical method could instead have produced the result

$$3x + 4 = 0 + (-4) + (-4)$$

or

$$3x + 4 + (-3) = 0 + (-4)$$

and so on.

Method  $\beta$  We can add *more formal content* by distinguishing the two sides of the equation. Rule  $\beta$  chooses *numbers* as method  $\alpha$  does, but always subtracts one number from one side of the equation, and subtracts the other number (which, of course, might or might not be the same number) from the other side of the equation. In short, method  $\beta$  does distinguish the two sides of the equation, but is unable to distinguish between the two input numbers.

Thus method  $\beta$  might respond to

$$5x + 12 = 0$$

with the "imitations"

$$5x + 12 + (-5) = 0 + (-12)$$

or

$$5x + 12 + (-12) = 0 + (-12)$$

or

$$5x + 12 + (-5) = 0 + (-5)$$

or

$$5x + 12 + (-12) = 0 + (-5) ,$$

but would NOT produce "imitations" such as

$$5x + 12 + (-12) + (-5) = 0 .$$

Method  $\gamma$  . Still more formal content can be added: method  $\gamma$  observes that in all of the illustrative examples, the *same* number is subtracted<sup>7</sup> from each side of the equation. [Method  $\gamma$  might arrive at this additional constraint by observing  $ax + b = 0$  , accepting  $a$  and  $b$  as input numbers *without* distinguishing between  $a$  and  $b$  (so that, in effect, method  $\gamma$  uses the set  $\{a, b\}$ , as input), then responding with  $c$  (where  $c$  is either  $-a$  or  $-b$ ); then for the selection of the next response, method  $\gamma$  uses the ordered pair  $(\{a, b\}, c)$  as the input. With this input record, it easily becomes apparent that the textbook's *second* response is always  $c$ . Note that this would be consistent with the hypothesis that one record that is maintained is a record of procedures used in the recent past.]

*As one develops an increasingly "meaningful" approach to imitation, it is quite remarkable to see where "meaning" is located; thusfar, such (small) amount of "meaning" as has been added has been of an entirely FORMAL nature. Nothing added thusfar would seem to be essentially semantic in the sense that one might have expected.*

<sup>7</sup>

For brevity, we do not here distinguish "subtracting  $N$ " from the equivalent operation of "adding the additive inverse of  $N$ ."

Method  $\delta$  . At this point we try to provide for some representations of "knowledge" of a more semantic nature.

In an actual class, whenever students subtract *different* numbers from the two sides of an equation, the teacher will (presumably) correct them, saying that "the *same* number must be subtracted from each side of the equation" - or words to that effect. Method  $\delta$  includes this explicit verbally-coded information.

Method  $\epsilon$ . The teacher, or the textbook, or the sequence of problems might go further, by introducing the concept of *truth set of an open sentence* [Davis, 1964, chapter 23], the idea of operations on equations that *do not change the truth set* [Davis, 1964, calls these "transform operations"], and the subgoal of *simplifying the equation without changing the truth set*. There may also be prior development of the concept of "undoing" certain operations (by using additive or multiplicative inverses) [Davis, 1964, Chapter 20]. We define method  $\epsilon$ , then, as one that "knows" of the sub-goal of "simplify the equation without changing the truth set" and which has a collection of procedures that are guaranteed NOT to change the truth set - e.g. "add the same number to each side of the equation" or "multiply each side of the equation by the same non-zero number."

We have NOT yet introduced any method for distinguishing the two original numbers - i.e., the  $a$  and  $b$  in

$$ax + b = 0 \quad .$$

We now do so:

Method  $\zeta$ : Distinguish the two original input numbers as "left" or "right" number. This will produce *correct* imitations for all problems of the form

$$3x + 5 = 0$$

but will produce *incorrect* imitations for all novel problems of the form

$$5 + 2x = 0 \quad ,$$

namely:

$$5 + 2x + (-2) = 0 + (-2) \quad ,$$

since the illustrative example in the textbook used the "right" number for subtraction.

Discussions with Herbert Simon, James Greeno, John Seely Brown, Marilyn Matz, Jill Larkin, and Lauren Resnick (see below) have pointed out the importance of *internal critics* - some capability which the student has *within himself* for detecting errors. Where "internal critics" are concerned, individual differences from one student to another are extreme; some students possess a very large repertoire of internal critics, and use them frequently, while other students seem content with "flagrantly wrong" answers where errors should be easily detectable.

"Internal critics" are probably of many different types, and (at least for adult experts) include items such as:

- i) The sum of an odd number of odd integers must yield an odd integer as an answer - otherwise there is an error;
- ii) Many forms of estimation - for example,

$$27 + 39 = 86$$

*must* be wrong, since  $27 < 30$ , and  $39 < 40$ , so  $27 + 39$  must be less than 70;

- iii) And, probably, checks against whether things "make sense". For example, for equations such as

$$5x + 3 = 0$$

it *MUST* be wrong to distinguish the two input numbers (5 and 3) as "left" or "right," *because one can reverse these roles without changing the problem:*

$$3 + 5x = 0$$

must be the same equation, thanks to the commutative law of addition, as

$$5x + 3 = 3 + 5x \quad .$$

Hence the "left" vs. "right" discrimination is failing to match whatever *true* discrimination may exist within the problem.

Method  $\eta$ :      Employ the *internal critic* just described, that would reject method  $\zeta$ 's "left" vs. "right" discrimination of input numbers.

One could easily hypothesize other "internal critics," such as:

Method  $\theta$ :      The additive inverse "undoes" addition. Hence it *may* make sense to write

$$5x + 3 + (-3) \quad ,$$

but it probably does NOT make sense to write

$$5x + 3 + (-5) \quad .$$

An important form of mathematical knowledge possessed by experts is knowledge of "canonical forms," such as

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

$$ax^2 + bx + c = 0$$

$$\frac{d^2 y}{dx^2} + f_1(x)y = f_2(x)$$

and so on. We add some canonical form" or "pattern matching" knowledge by hypothesizing:

Method  $\iota$  :      Identify the two input numbers by matching the equation in the illustrative example

$$3x + 4 = 0$$

against the canonical form

$$ax + b = 0 \quad .$$

Note: Method  $\iota$  guarantees a correct identification of the two input numbers, provided the equation is of the form

$$ax + b = 0 \quad .$$

Otherwise method  $\iota$  is inapplicable.

To deal with equations not of the form

$$ax + b = 0,$$

we hypothesize a more complex set of procedures which we do not presently analyze: procedures capable of re-writing some other equations (we leave the specification of just *which* equations open for the present) in the form  $ax + b = 0$ . For example, this method presumably can call upon the commutative law of addition to rewrite

$$7 + 3x = 0$$

in the form

$$3x + 7 = 0.$$

Call this incompletely-defined capability method  $\kappa$ .

But, as Anne (or the computer program) studies illustrative examples in the textbook, they may make observations of a different form: they may observe the *effect* of textbook operations. For example, the first textbook operation *has the effect of producing an expression of this form*

$$ax + b + (-b) \quad .$$

Call the capability of making this observation method  $\lambda$ .

*Note:* method  $\delta$ , with its "meaningful" knowledge that one is supposed to subtract the *same* number from each side of the equation, could function as an *internal critic*. We illustrate this, below, in Diagram IV.

As Anne, or the computer, tries to learn how to "imitate" successfully the illustrative examples in Stein and Crabill, one can attempt to chart their progress with diagrams wherein *nodes* (labeled with English letters)

represent definite patterns of imitation, and the directed *links* connecting the nodes represent what has been "learned" by using the various methods listed above. Links are labelled by the Greek letter that identifies the method in question (thus, a link labelled " $\beta$ " refers to progress by using method  $\beta$ , and so on). Links drawn downward *to the left* represent learning by an increment of *formal* knowledge, those drawn downward *to the right* represent increments of *semantic* knowledge. Here "semantic knowledge" or "meaning" is taken to be knowledge of *constraints* ( $\delta$ ), *goals* ( $\epsilon$ ), *canonical forms* (considered as a special class of *assimilation paradigms*) ( $\iota$ ), and *organized collections of procedures (or frames) called up from previous learning* ( $\kappa$ ).

"Internal critics" are shown *canceling* some increment as unacceptable, and consequently requiring a return to a previous node ( $\eta$ ;  $\delta$  when it functions as a "critic").

In Diagram I we see a computer (or Anne) proceeding at first with a minimum of semantic knowledge, and a maximum of "blind imitation," with semantic knowledge added only in the later stages. The "blind imitation" processes  $\delta$ ,  $\beta$ , etc., do NOT, of course, lead to rapid convergence toward correct methods of solving linear equations.

Among matters *not* dealt with at this level of analysis are the mechanisms by which the obtaining of incorrect results produces a search for alternative methods for future use.

In Diagram II we ask: how would things be different for a student who *began* his study of Stein and Crabill's First Example *already familiar with linear functions* (which is supposed to be the case for the students at University High School)? Hence, method  $\iota$  is available from the outset, and confusing the two input numbers - or failing to distinguish between them correctly - should not occur. Given the semantic knowledge  $\epsilon$  and  $\kappa$ , (which, again, University High School students are supposed to possess at this stage),

start: observe illustrative example  
in text book

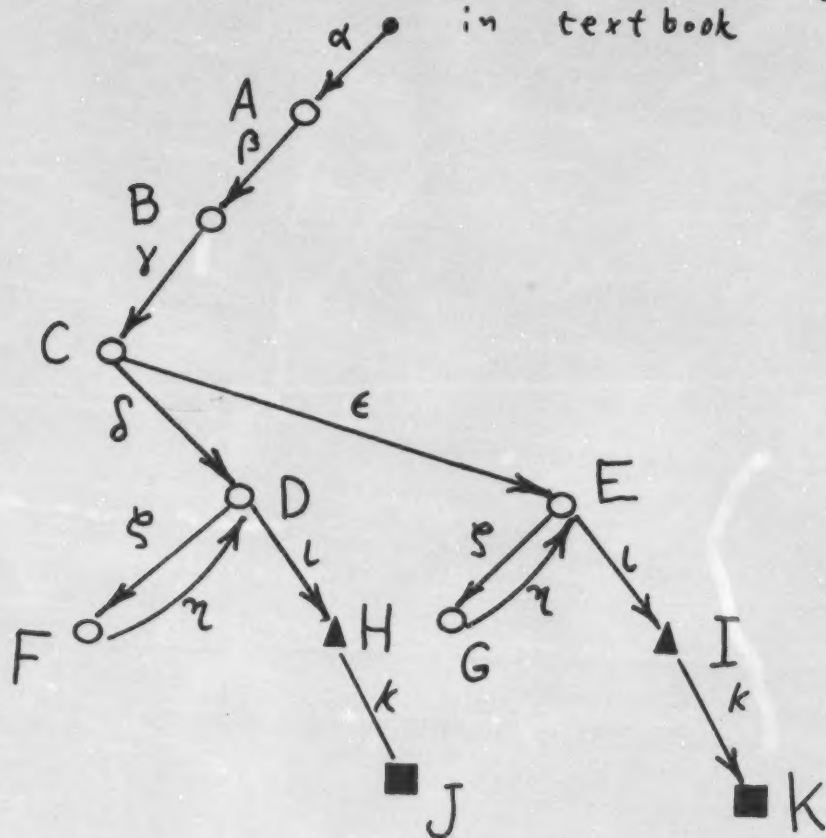


Diagram I

- code:
- indicates a faulty imitation
  - ▲ indicates a correct solution for all equations of the form  $ax+b=0$ .
  - indicates a correct solution for all equations of the forms  $ax+b=0$  or  $a+bx=0$ .

Start: observe illustrative example  
in text book

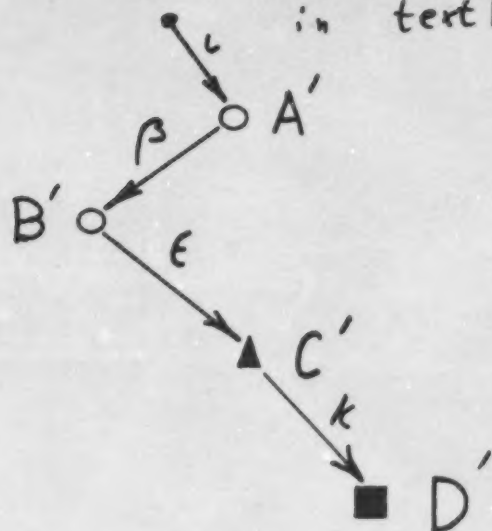


Diagram II

Start

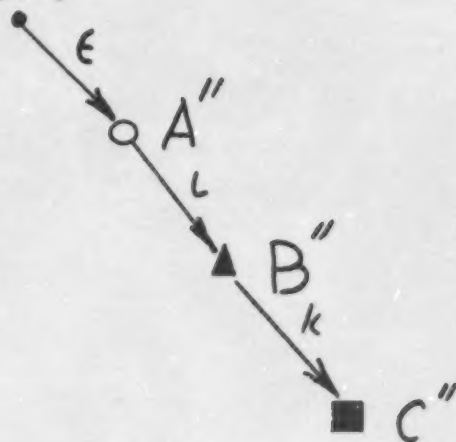


Diagram III

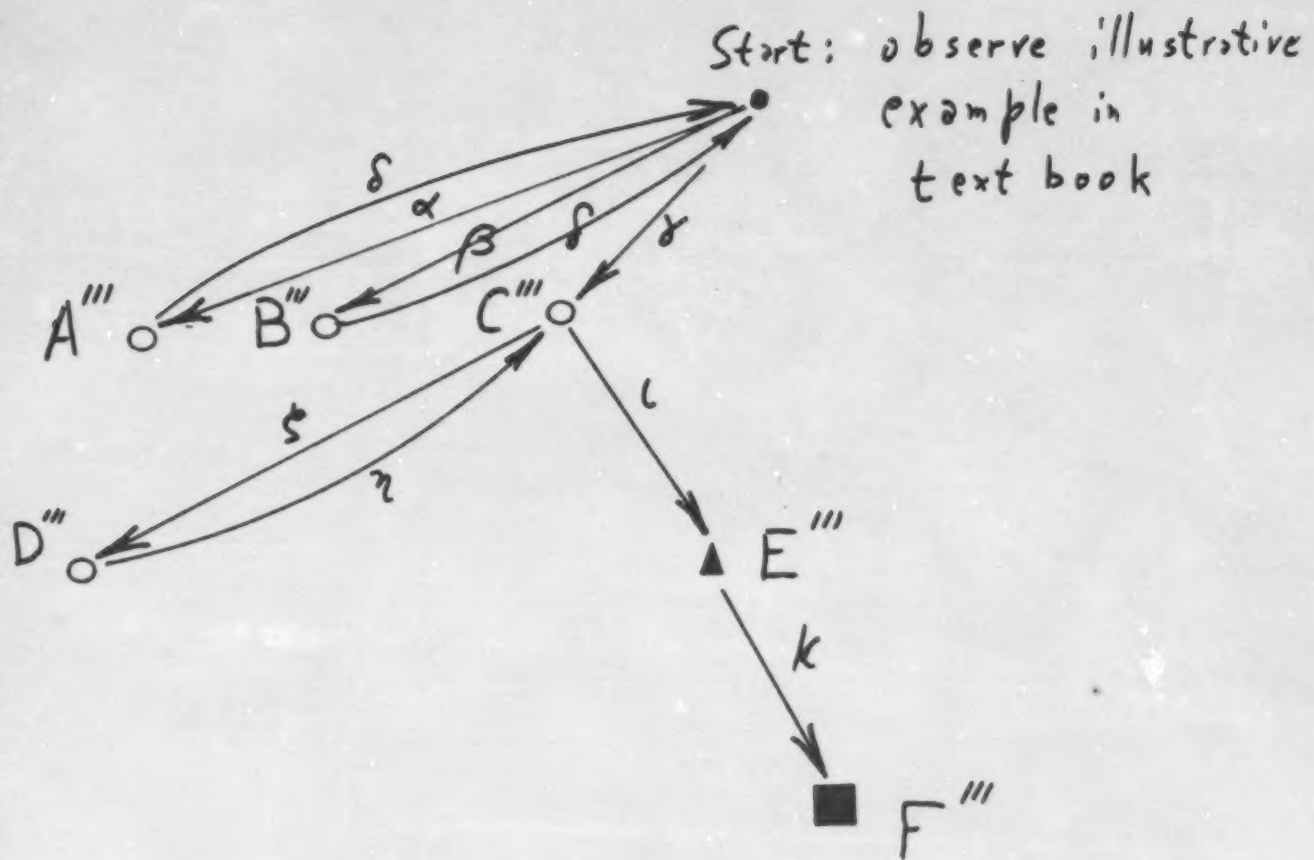


Diagram IV

convergence to a correct method of solution should be fairly rapid.

In Diagram III we ask a similar question, but assume that the student has even more semantic knowledge available, and makes effective use of it.

Finally, returning to our student John, who seems to get things right very early on in novel situations, we ask: what would happen if a student had a large collection of *internal critics*, and *made very extensive use of them*? (From remarks that John often made, there is reason to suspect that this is descriptive of him.) Thus, he tries  $\alpha$ , but critic  $\delta$  leads to immediate rejection, and a return to the starting node; he tries  $\beta$ , but critic  $\delta$  again rejects this. Method  $\gamma$  satisfies critic  $\delta$ , leading to Node C<sup>'''</sup>. He uses method  $\zeta$  (a really "mindless" method, but one found often in C. Resnick's study of a PLATO lesson cf. Resnick, 1975), but critic  $\eta$  immediately rejects it, returning to Node C<sup>'''</sup>. Methods  $\iota$  and  $\kappa$  then lead to correct solutions.

Does this describe John? The answer is that it is too early to say; work on this representation of learning is just beginning. It does, however, seem more promising than any previous analysis that has been attempted; at least it matches certain known attributes of John's behavior, such as his volunteered remarks indicating reference to "internal critics," and it matches his known speed in arriving at correct methods from inspecting illustrative examples.

#### 4. Ricky's Method

We have seen earlier how Ricky, in grade 7, prior to formal instruction in solving equations, solved

$$\frac{3}{x} = \frac{6}{3x - 5}$$

by saying [paraphrased]: "On the right hand side the numerator has been doubled. So the denominator must also have been doubled. But going from  $x$  to  $3x$  means adding one  $x$  too many - this must be exactly offset by subtracting

5. Hence  $x$  must be 5."

We are now systematically trying to identify this stage of analyzing

algebraic equations among 7th grade students who have NOT had formal instruction in solving equations. One importance of this stage is that, when present, it is a positive indication that the student is able to use the "x" notation correctly as the name for a presently-unknown number. E.g., to use his method, Ricky *had* to know that the "x" on the left-hand side of the equation, and the "x" on the right-hand side, *both named the same number*.

### 5. *The Clement-Kaput Phenomenon*

James Kaput, of Southern Massachusetts University, and John Clement, of the University of Massachusetts, report an interesting phenomenon - a common and persistent error among relatively sophisticated students in translating the sentence "There are 6 students for each professor," into the algebraic expression

$$6 S = P, \quad (1)$$

despite the fact that equation (1) should immediately be seen to be in error (if one applies any "internal critics" relating to the relative size of the number S (the number of students) and the number P (the number of professors), and the effect of multiplying a positive integer by 6.) [Clement and Kaput, 1979].

In our own next work, we shall attempt to use task-based interviews to detect and analyze this phenomenon in beginning algebra students. It is of considerable importance because of its implications for information processing by algebra students.

### B. *Use of the Rhetoric*

Part of the "product" of the present study has been the development of a rhetoric for the discussion and analysis of information processing activities by students who are learning mathematics. In this section we show how this rhetoric can be applied to empirical results obtained by other investigators, in reasonable detail, in the case of the Friend phenomenon, and more sketchily in the case of two phenomena that lie outside our original field of investigation.

1. *The Friend Phenomenon*

Naively, one would assume that the addition problem

$$\begin{array}{r} 121 \\ 32 \\ \hline 23 \end{array} \quad (2)$$

would be easier (more students would answer correctly) than either the problem

$$\begin{array}{r} 121 \\ 132 \\ \hline 25 \end{array} \quad (3)$$

(because it requires one more addition), or the problem

$$\begin{array}{r} 121 \\ 32 \\ \hline 63 \end{array} , \quad (4)$$

which involves a "carry" from one column to another. Jamesine Friend, of Stanford University, has reported that this is not the case, but, on the contrary, problems (3) and (4) are easier than problem (2) [Friend, 1979].

This is a genuinely surprising result. Analysis reveals that the greater difficulty of problem (2) lies in the left-most column. It can be analyzed as follows: the primary-grade addition frame (discussed above) requires two inputs. To get a second input, students resort to various incorrect procedures, such as adding the two numbers in circles!

$$\begin{array}{r} 121 \\ 32 \\ \hline 23 \end{array} ,$$

hence getting the answer 376, instead of the correct answer 176.

This surprising result becomes entirely *unsurprising* when one analyzes the information processing activities in terms of the primary-grade arithmetic frames, such as:

|            |                                     |
|------------|-------------------------------------|
| add        | This requires two inputs; get them! |
|            | one input: _____ (A)                |
|            | other input: _____ (B)              |
| Return A+B |                                     |

## 2. *The Place-vs.-Response Controversy in T-Mazes.*

We have made reference, earlier, to studies of rats running mazes, and especially T-mazes. Prior to (about) 1957, there was a controversy over whether rats learned a sequence of *decisions* ("left, then right, then right again, ..."etc.) in learning complex mazes, or whether they learned a definite *place* as the goal ("go over toward the cages where the other rats are..."). In the July, 1957 issue of *The Psychological Review*, Restle presented a discussion, based on the theory of cue discrimination, that indicated that the controversy had not been formulated properly. What the rat learned depended upon the *proportion* of relevant cues; the totality of cues included visual, auditory, olfactory, tactual, and kinesthetic stimuli, and might be associated with the maze itself, or with the room in which the maze was located. Cues of different sorts tended to carry different weights - "where the cages with the other rats are" was an extra-maze cue of great influence; the presence of small contrasts in level of illumination was NOT very influential. Within a given experimental set-up, one collection of cues would relate to "responses", another collection of cues would relate to "place" (depending upon how maze was or was not rotated, which outcomes were rewarded, etc.).

One part of the early maze theory, as we have remarked earlier, matches closely with the conceptualization developed in the present study: the idea that *VMS sequences* can "solidify" into *integrated sequences* because of the operation of some kind of "look-ahead" capability. This is a precise match with the "kinesthetic" theory [Restle, p. 219; Small, 1901; Carr and Watson, 1908] of the early twentieth century discussion, that each response would provide kinesthetic cues to trigger the sequentially following response, thus producing an "automatic" performance. This is precisely our transition from VMS sequence to integrated sequence, except of course that we are interested in general information processing procedures, and not any specific sensory modality.

But the "place-vs.-response" controversy has at least two other points of

contact with our work.

First, there is the question of "boot-strapping" cues<sup>9</sup>; one of the present authors spent two weeks on the campus of the University of California at Santa Cruz. Daily activities required finding one's way around the campus, which was at first very confusing. Gradually, however, key decision points became recognizable. Presumably something was being "learned" - indeed, presumably some cues (mostly visual) were being correlated with correct decisions. The mystery, however, lies here. Suppose in fact that "every spot looked exactly like every other spot," which is the sensation one had, more or less, at the beginning. In that case there would be no visual cues that could serve to distinguish one spot from another, one collection of input data from another. *On the "visual cue" side of the correlation, there would be no input data to correlate.*

One must, then, assume some subtle differences (there were, of course, many) that at first escape notice, but gradually are selected and, in the classical sense, "conditioned" [as in Restle, p. 218]. It should be observed, then, that future learning depends upon the success of this bootstrapping operation, this lifting of some stimuli above a random "noise" level, and hence future learning *depends upon past achievements in learning methods for discriminating such stimuli.* If, for example, you can recognize *Arctostaphylos tomentosa*, and if you can thus determine that there is only one *A. tomentosa*, which happens to indicate a place where you need to turn left, then you're "learning the maze" quite nicely; but if, on the other hand, you can't distinguish *Arctostaphylos tomentosa* from *Arctostaphylos pungens*, or even from juniper, white pine, or eucalyptus, then these cues are "invisible" to you. What happens then? Does some primitive olfactory cue (or

---

9 The "boot-strapping" analysis was suggested by John Seely Brown.

something of the sort) penetrate into consciousness, and become an attribute of an increasingly "recognizable" spot on the campus? Even though you've never encountered it elsewhere? Even though you have no name for it? Even though you find it entirely "weird" and have "no idea what it is?"

If you already possess an adequate set of descriptors, then "learning by correlation" in the maze (or VMS) sense is a comprehensible process. But if you do NOT possess an adequate set of descriptors, what then? The John Seely Brown suggestion of "bootstrapping," and the Restle analysis of maze learning, both indicate the same answer: *some* cues must be dredged up from the "bloomin, buzzin' confusion," given some sort of identity or label (not necessarily verbal), and put to work to extend your list of descriptors.

The matter of having appropriate descriptors is clearly crucial to effective functioning, in mathematics or elsewhere. One can bring this closer to the present study by considering the famous *Logic Theorist*, a computer program for the RAND Corporation's JOHNNIAC computer, that could create original proofs for theorems in symbolic logic [cf. Newell, Shaw, and Simon, 1958]. The logic theorist ("LT") can attempt to prove a theorem by generating possibly useful formulas using replacement of variables. Infinitely many different formulas *could* be generated in this way. Hence, the "straightforward" method of generating and generating and generating, and *hoping* to get the desired theorem as a result, is unfeasible. LT uses descriptors to describe the goal theorem, and uses these same descriptors to describe the axioms, and avoids using axioms whose descriptions do not match.

But, of course, one has to have the set of descriptors.

Beginning mathematics students do not ordinarily possess a useful set of descriptors. [In the LT system, one description refers to "an expression that has a single variable, one argument place on its left side, and two argument places on its right side." Examples that satisfy this description include

$P = (pVp)$

and

$r = (r \quad r) \quad .]$

This observation may have pedagogical implications: perhaps far more effort ought to be expended in teaching beginning students a useful set of descriptors capable of describing mathematical formulas, concepts, and other entities.

Finally, there is an even more immediate link with the present study. In fact, as Restle's analysis suggests, the real interest in the "place-vs.-response" controversy cannot truly have meant either "place" or "response." The real question is the extent of cognitive knowledge of geometry that exists in the mind of the rat. To what extent does the rat possess a 2-dimensional map?

Turning again to personal experience, one of the authors once had three different "maps" of greater Boston stored in memory: a subway map, an automobile map, and a pedestrian map. At first, he knew almost no relations between these different maps. When a rat has learned, say, a tunnel maze, does he have any picture of the two-dimensional space he is traversing? Simon has presented some convincing arguments that humans do NOT have genuinely geometrical "pictures" in their minds, so presumably rats do not, either. But a rat MIGHT have in his repertoire some heuristics for adjusting paths - so that, for example, when one path is blocked the rat might have a sophisticated strategy for finding an unblocked equivalent path. Or, if the rat has not developed such heuristics from his general experience with movement and space, perhaps he could be taught some heuristics for finding replacement paths.

### 3. *Ivan's Brother's Brother*

Zankov [1977] reports on one of the largest educational R and D ventures in history, a program that redirected the education of primary-grade children in the Soviet Union. In reference to psychological foundations, Zankov [p.17] speaks of

a child who, from "practical experience" knows that, say, Igor is his brother. This is knowledge of a very practical, immediate, experiential, "concrete", and complete type. Presumably the child has one (or many) *frames* dealing with Igor, that include the names that proclaim that Igor is "my brother." Perhaps the child can go one stage further, and may know the concept of *brother*, not attached to the specific instance of Igor. This, presumably, is a frame somewhat like this:

x is *brother* of y

[x must be male]

- |                           |     |
|---------------------------|-----|
| 1) determine parents of x | A,B |
| 2) determine parents of y | C,D |

If {A,B}, {C,D}  $\neq \emptyset$  ,

x is brother of y

[There are details dealing with "half-brother,"  
"step-brother", etc.]

[A fundamental question is how this *abstract* frame for "brother" can have been created from *specific instances* frames for Igor, etc.]

But Zankov goes still further, and asks whether the child can deal with the concept of *Ivan's brother's brother*. In information processing terms, what is involved? We give two brief outline descriptions, which are presumably equivalent to one another:

*First description:* To deal with this last problem, the child needs some procedure for "shifting point of view," roughly like this:

To shift point of view

- 1) record input ["Ivan"]  
 $P \leftarrow$  Ivan  
 find *brother* of P

- 2) record input ["brother of Ivan"]  
 $P \leftarrow$  brother of Ivan  
 find *brother* of P

## 3) retrace records.

*Second description.* An alternative, but possibly equivalent description, can be couched in the language of LISP [Friedman, 1974] or LOGO [Feurzeig et al., 1972], languages that have a built-in recursion capability. To illustrate [again, roughly] one could define the factorial function recursively somewhat like this:

Define:  $N!$

- 1) If  $N=1$ , return 1
- 2) Find  $(N-1)!$  and return the product  $N \times (N-1)!$

Watch this recursive program compute 4!

```

N=4
Find 4!
  N=1 is false; proceed to step 2
  record N=4, find 3!
    Find 3!
      N=1 is false; proceed to step 2
      record N=3, find 2!
        Find 2!
          N=1 is false; proceed to step 2.
          record N=2, find 1!
            Find 1!
              N=1 is true; return 1.
            Multiply 2 × 1; return 2 × 1.
          Multiply 3 × (2 × 1); return 3 × (2 × 1).
        Multiply 4 × [3 × (2 × 1)]; return 4 × [3 × (2 × 1)].

```

It should be clear that the program has worked correctly.

One can use this *recursion* feature to enable the child to cope with *Ivan's brother's brother*, via a program roughly like this:

*P's brother*

- 1) find P's parents
- 2) find male child of P's parents, not equal to P

*Ivan's brother's brother*

- 1)  $P \leftarrow \text{Ivan}$
- 2) find brother of (brother of P)

The greater maturity of the child who can cope with "Ivan's brother's brother" presumably consists either of some relatively "concrete" frame for "shifting point of view", or else of some quite general capability to use LOGO-like recursion.

[Incidentally, we have in our files one case that relates to precisely this usage of language. Alex, aged 3 years, 6 months, has answered the telephone, and comes to her father, saying: "My mommy... *her* mommy...is on the 'phone.'"]

### C. *Future Work (Theory)*

Compared to the size of the region to be explored - a very large number of students studying mathematics with many thousands of teachers and hundreds of different curricula, for perhaps twelve years (or more) per student - this has obviously been a very small project. We feel confident that it is a useful beginning - but it is only a beginning. There is a need to build up a far larger collection of relevant behaviors - recorded, as much as possible, on videotape. There is an evident need to sharpen and extend the conceptualization or "rhetoric" for discussing and analyzing mathematical behavior. And there is an especially promising direction that needs to be pursued: the identification of especially illuminating phenomena, such as the Friend phenomenon and Clement-Kaput phenomenon.

From our work thusfar, we have been struck by the need to gain a deeper understanding of six fundamental information-processing capabilities:

#### 1. *How "new" ideas are built up on a foundation of earlier ideas.*

We have seen several aspects of this. The "binary switches" considered in Davis, Jockusch, and McKnight, (1978), are one example. The building up of the ideas of *algebraic addition* and *algebraic subtraction* on a foundation of the primary-grade notions of "adding" and "subtracting" (including symmetric subtraction) is an especially interesting story (cf. Davis and McKnight, 1979A;

Davis, McKnight, Parker, and Elrick, 1979). Similar "developmental" themes are associated with most of the fundamental concepts in mathematics.

Figuratively speaking, mathematical ideas, like many ancient cities, are built atop the ruins of their predecessors, and careful cognitive studies often have something of the character of archaeology, turning up traces of earlier versions of an idea. (For example, the earliest meaning of the equals sign = is *not* symmetric:  $3 + 5 = 8$  is meaningful, but  $8 = \underline{\hspace{1cm}}$  is not; this early version is strikingly similar to the use of the "=" key on a hand-held calculator. This early meaning is roughly equivalent to saying: "and the standard numeral for this is           ". Later meanings depend upon the specific curriculum a child has experienced; if, for example, he has worked with balance beams, in the British style, then

$$3 + 2 = 5$$

$$5 = 3 + 2$$

$$3 + 2 = 2 + 3$$

are all (perhaps among other things) statements about where washers are hanging on a balance beam. Statements like

$$3 + 2 = 2 + 3$$

come to be seen as descriptions of "how numbers work." At a more mature level, *one* meaning of equality

$$A = B$$

is that "A names the same thing that B names.")

2. *Descriptors.* We have already seen the fundamental role of descriptors: descriptors of symbol strings, descriptors of functions, descriptors of geometric proofs, etc. In those (typical) cases where descriptors are not taught explicitly, *where do descriptors come from?* [A partial answer is suggested by the way a *simple closed curve* can be defined by the *act* of traversing it, observing

whether one encounters points of intersection. *Some new descriptors are created by carrying out some process and observing the result of this process.*

3. *Metaphors.* Few, if any, ideas are more central than the idea of metaphor. When are two "different" things capable of being, in a limited sense, identified? [E.g., one must be able to recognize

$$e^{2t} - 53^t + 6 = 0$$

as (among other things) a "quadratic equation."]

4. *Selecting Assimilation Paradigms, Mapping Input Data into the Receptor Slots("Variables"), and Judging Correctness of the Result.* This, of course, is absolutely fundamental.

5. *Meta-Language or Planning Language.* When the task of finding the distance from a point to a plane is first encountered as a novel problem, an original solution method must be invented. This is easily done, by selecting and sequencing appropriate sub-procedures, *provided one possesses appropriate descriptors for these sub-procedures* (or an appropriate planning language).

6. *Dealing with Knowledge Coded in Different Representations.*

Presumably some mathematical knowledge is coded algorithmically, some as formulas, some as natural language statements, and some geometrically. How are these different kinds of knowledge brought together and used in effective combination? [Cf., for example, the task of learning to solve linear equations by studying the illustrative examples presented in Stein and Crabill.]

## VII. *Educational Science Today*

### A. *Points of View*

Kuhn [1962], Polanyi [1958], and others have emphasized that one of the most important consequences of scientific activity takes the form of a *change in point of view*. [Cf., e.g., Polanyi on the definition of *isotope*.]

In the late nineteen fifties the typical point of view in mathematics education research was that one could not study *ideas*, which were unknowable mysteries that *might* be present in the minds of subjects. To be "scientific," one must deal only with that which was observable. That this doctrine would have turned physics into the study of meters, dials, needles, and flashing lights was, of course, overlooked. Critics quipped that man, "having lost his soul, and his freedom, had now lost his mind." Mathematicians found the typical point of view unacceptable because they believed that mathematics consisted of ideas in the minds of mathematically-proficient human beings (and, up to a point, in the minds of birds, kittens, etc.).

The "hard data" often used in the quest for "being scientific" often consisted of numbers of students making the various choices on a multiple-choice test. Perhaps no better commentary on this approach exists than the remarks of Krutetskii:

"...The overwhelming majority of tests do not rest on solid scientific grounds, which makes the psychological meaning of test studies unclear...

A basic defect in test research is the bare statistical approach to the study and evaluation of abilities - the fetishistic mathematical treatment of test results, *with a complete absence of interest in studying the solution process itself* [italics added]. ...Failing to reveal the mental processes leading to a given result, these tests do not provide a complete picture of the phenomenon under scrutiny. ...

As an illustration of the aforementioned deficiency of test studies, we can cite the research of J. Murray, devoted to an analysis of geometrical ability and published in 1949. The author, using a testing technique that completely ignores the process of solving test problems, found a correlation between success in mastering geometry and particular abilities: reasoning ability, the ability to "grasp" spatial relations, and so on. The same kind of research is represented by the more recent investigations of mathematically gifted adolescents by W.A. Kennedy and others, the research of E. Locke on the correlation between indexes of gifted science students' intellectual abilities and personality traits obtained in the classroom and their out-of-class achievement, and many more. It is hard to understand how theory or practice can be enriched by, for instance, the research of Kennedy, who computed, for 130 mathematically gifted adolescents, their scores on different kinds of tests and studied the correlation between them,

finding that in some cases it was significant and in others not. The process of solution did not interest the investigator. But what rich material could be provided by a study of the process of mathematical thinking in 130 mathematically able adolescents! [Krutetskii, 1976, pp. 13-14]

Krutetskii quotes one researcher, whose attitude he thinks quite wrong, as having said "only external manifestations constitute reality;...one should not try to persuade oneself that something deeper is being studied." [Krutetskii, op. cit., p. 13]

This concern for "product" only, with a deliberate and systematic avoidance of "process", this narrowing down the notion of product to (in many cases) nothing more than selecting one from among four possible choices, and this determined renunciation of any interest in thought processes themselves, severely limited the possibilities for research and development in mathematics education two decades ago.

One might have thought that the appearance of papers such as the Newell, Shaw, and Simon paper (op. cit.) in 1958 would mark the beginning of a new era, and perhaps it did, but if so the period of gestation was extremely long. It is only now, in 1979, that an information-processing analysis of mathematical behavior is becoming at all common.

#### B. *Communications*

The recent rapid growth of "process" studies of mathematical thought, and of cognitive science or AI conceptualizations for the interpretation of the results, clearly owes much to the fundamental contributions and continuing effective leadership of Herbert Simon, Marvin Minsky, and Seymour Papert. Certain forums are developing for the exchange of information at a profound and current level, especially the (approximately) annual invitational conferences sponsored by James Greeno, Robert Glaser and the Learning Research and Development Center of the University of Pittsburgh. These are the most valuable scientific meetings

that the present investigators have ever been privileged to attend. The contributions of Herbert Simon, Jill Larkin, John Seely Brown, Marilyn Matz, and others contribute importantly to this result. We hope that the reports of the present project have made a useful addition to this extremely valuable dialogue.

The San Francisco meeting of the American Educational Research Association also provided an important forum for discussions of the cognitive science approach to the study of mathematics learning. Particularly important were contributions by Guy Groen, Ira Goldstein, Jamesine Friend, John Seely Brown, Marilyn Matz, and Edward J. Davis, among numerous others.

Several journals now present this kind of study, notably the *Journal of Children's Mathematical Behavior* and *Cognitive Science*. Occasional studies of this type appear in the *European Journal of Science Education* (and, of course, the seminal Newell, Shaw, and Simon study appeared in *The Psychological Review*).

There is also a growing tendency to coordinate related studies, for example by using the same task problems in different studies. We hope in this way to be able to relate some of our own future work to work by Clayton Lewis and John Bernard, at the University of Texas, and to the work by the Simon group on learning to solve linear equations.

### C. *Relation to Other Approaches*

Other important approaches to testing, and to the forms of learning and performance that underly testing, include the work of Piaget and his collaborators, Gagné, Zacharias and his collaborators, and Krutetskii.

*Piaget.* The Piaget group undertook to create an evaluation program for the British Nuffield Mathematics Project, under Geoffrey Matthews. Neither the British nor the Geneva side of this collaboration had confidence in traditional tests; as an alternative they created "check-ups." These check-ups had much in common with our use of task-based interviews, the most conspicuous differences being that we study older children, our tasks are more clearly school-related, and

we make less use of special physical materials. Our *analyses*, however, are entirely different (as should be clear from the preceding discussion), and we do not recognize *stages* in the Piagetian sense. While we do recognize that new cognitive structures are built up, and that (in this sense) important developmental changes do occur, we always attempt to relate these changes to highly specific information-processing capabilities. We believe that even very familiar Piagetian tasks take on a new meaning when analyzed in terms of our conceptual apparatus.

*Gagné*. The work of Robert Gagné in this general area is well-known [cf. Gagné, 1967; Gagné, 1962]. It resembles our work in some regards, but differs drastically in other ways.

First, Gagné's work *resembles* ours in a focus on *general solution processes* and NOT on specific responses to specific stimuli. For example, Gagné writes:

The growing interest in autoinstructional devices and their component learning programs has had the effect of focusing attention on what may be called "productive learning." By this phrase is meant the kind of change in human behavior which permits the individual to perform successfully on an entire *class* of specific tasks, rather than simply on one member of the class. Self-instructional programs are designed to ensure the acquisition of capabilities of performing classes of tasks implied by names like "binary numbers," "musical notation," and "solving linear equations," rather than tasks requiring the reproduction of particular responses. [Gagné, 1962, p. 355]

While we and Gagné both deal with methods for solving a *class* of problems, our analyses are entirely different. We look for students developing certain "procedures," "sub-procedures," and so on. Thus, when a sixth grade girl "solved"

$$.3 + .4 =$$

$$3. + 4. =$$

$$.3 + 4. =$$

by writing

$$.3 + .4 = .7$$

$$3. + 4. = 7.$$

and

$$.3 + 4. = .7. ,$$

using the invented symbol

$$.7. ,$$

we look for *procedures* (again, in an A.I. sense) for locating the decimal point, and (again, borrowing from A.I.) we look for relations between different data bases (akin to Minsky's "frames"), as in asking questions about whether .7. is larger than 6, or smaller than one, and so on.

Moreover, as John Seely Brown has pointed out, once one chooses to use "computer program" metaphors to describe human information processing, it is neither necessary nor natural to assume similar performance "on an entire class of tasks," because procedures typically have bugs, and bugs - though conspicuously systematic - tend to treat different cases in different ways. [A bug, for example, might involve the faulty "addition fact" that  $8 + 5 = 12$ , or might fail only in the case, say, of the Friend phenomenon.]

As a second difference, to some extent a consequence of the first, where Gagné sees a class of tasks in a somewhat gestalt form, we (as necessitated by our analysis in terms of "procedures") look *within* each task, *at an extremely minute level of detail*.

The third difference - an essential one - follows from the first two. Gagné organizes his classes of tasks into a strict hierarchy:

There are no instances, for example, of an individual who is able to perform what has been identified as a "higher level" learning set, and who then shows himself to be unable to perform a "lower-level" learning set related to it. [Gagné, 1962, p. 360]

Our preliminary studies have produced results entirely at variance with this. The "procedures" which students acquire are NOT capable of being arranged in such a linear sequence, as any number of examples make clear. Indeed, Mat was able to multiply

and to get the correct answer, because the specific cues in the situation elicited a correct performance algorithm; but Mat could not be relied upon in such matters. Given slightly different cues, he produced different "answers," and was not disturbed by the fact that the same problem seemed to lead to different answers on different occasions. Gagné's curriculum is based on classes of tasks, organized on a strictly linear basis. Our analysis of cognitive growth is based on the acquisition of "procedures" or "algorithms" (in an A.I. sense), and is far from linear. While the Gagné organization has a convenient simplicity, especially for automatic routing of students through an individualized curriculum, we believe that our conceptualization offers a far better fit to the actual phenomena of a student's conceptual growth.

Finally, we infer that Gagné is NOT interested in "process" in the sense that we use the word. His "process" occasionally resembles more nearly what we would call "product."

*Krutetskii.* We feel very sympathetic to Krutetskii's approach (cf., e.g., Krutetskii, 1976), and indeed believe that the similarity between his work and our own is (somewhat paradoxically) greater and deeper than it seems to be at a superficial level. Of course, the actual metaphors or rhetoric that Krutetskii uses to describe thought processes differ markedly from those that we employ, but we believe, for any given data, his analysis tends to bear a close resemblance to our own.

*Zacharias.* Jerrold Zacharias and his colleagues at E.D.C. are carrying out an ambitious program aimed at creating new approaches to testing, and also aimed at creating acceptance for these new methods. In nearly all of their basic assumptions they are in harmony with our work, but their goal is more practical and immediate, in terms of producing effective improvements in testing programs that should begin to make a difference in the lives of U.S. students in the fairly near future. They also have a special focus on culturally-different children, and especially on low-SES urban minority children. We have acquired the highest regard for many aspects of their work, but we have not thusfar been able to build the "bridges" that would relate it effectively to our own

### VIII. *Details of Specific Activities*

The project became active October 1, 1976, and terminates September 30, 1979, the preparation of this report being one of its final activities.

The Director and Principal Investigator has been Professor Robert B. Davis, assisted by Dr. Curtis C. McKnight. Virginia Bradford, a teacher at a local secondary school, and Nancy MacDonald and Douglas Elrick, teachers at two local elementary schools, have played important roles in the work. Seven graduate students (Hiroko Kataoka, Steve Young, Vance Christiaan, Elka Block, Myriam Steinback, Mary Ann McClary, and Josephine Vamboi) have worked on the project in different capacities, with Steve Young playing a major role as a cooperating teacher, and Myriam Steinback working on certain research aspects. Elizabeth Jockusch was also an important contributor, as was Sharon Dugdale.

*Presentations of results* were made at a symposium of educators in Chicago, Illinois, on October 27 and 28, 1976; at Education Development Center in Newton, Massachusetts in November, 1976; at West Chester State College (Pennsylvania) in April, 1977; at Michigan State University in May, 1977; at the University of California (Berkeley) in October, 1977; at San Francisco State University in October, 1977; at the California Mathematics Council in October, 1977; at the Teachers Center for San Diego County, October, 1977; at the National Council of Teachers of Mathematics (NCTM) meeting in Indianapolis in November, 1977; at an N.I.E.-sponsored seminar in Washington, D.C., in January, 1978; at the A.E.R.A. meeting in Toronto, Canada, in March, 1978; at the first of the Greeno sessions at the LRDC, University of Pittsburgh, in September 1978; at the N.I.E.-sponsored conference on testing in August, 1978; at the University of Missouri (St. Louis) in October, 1978; at the Illinois Council of Teachers of Mathematics, October, 1978; at the Arkansas Council of Teachers of Mathematics, November, 1978; at the Educational Services Center

in Davenport, Iowa, November, 1978; at Teachers College, Columbia University, February, 1979; at four major presentations at the San Francisco meeting of the American Educational Research Association in April, 1979; at the International Study Group for Mathematics Education meeting in Boston, Massachusetts, April, 1979; at the NCTM annual meeting, Boston, Massachusetts, April, 1979; at California Polytechnic University, San Luis Obispo, May, 1979; at a National Science Foundation workshop at Wesleyan University, Middletown, Connecticut, July, 1979; and at the extremely important second Greeno meeting at the Pittsburgh LRDC Center, July, 1979. In addition, aspects of the work served as inputs at several other meetings convened by N.I.E., by N.S.F., and by E.T.S.

*Publications:* Publications (or unpublished reports, as indicated) resulting from the work of this project include:

Davis, Robert B.

"Representing knowledge about mathematics for computer-aided teaching, part I. - Educational applications of conceptualizations from artificial intelligence," in E.W. Elcock and Donald Michie, eds., *Machine intelligence 8: Machine representations of knowledge*. Ellis Horwood Ltd. (dist. by Halstead Press, John Wiley & Sons, Inc., New York), 1977, pp. 363-386.

Davis, Robert B., Sharon Dugdale, David Kibbey, & Charles Weaver.

"Representing knowledge about mathematics for computer-aided teaching, part II. - The diversity of roles that a computer can play in assisting learning," in E.W. Elcock and Donald Michie, eds., *Machine intelligence 8: Machine representations of knowledge*. Ellis Horwood Ltd. (dist. by Halstead Press, John Wiley & Sons, Inc., New York), 1977, pp. 387-421.

Davis, Robert B., Elizabeth Jockusch, Curtis McKnight, and Stephen Young.

"Observer notes on student and teacher behavior." University of Illinois Curriculum Laboratory (unpublished), 1977.

Davis, Robert B.

"Curricular implications for research on teaching mathematics." Paper presented at Michigan State University, East Lansing, Michigan (Institute for Research on Teaching). May 1977.

Davis, Robert B., Curtis McKnight, and Elizabeth Jockusch.

"Progress report on NIE grant no. HEW NIE/G-76-0085. June 15, 1977. Report submitted to the Nat'l Institute of Education.

- Davis, Robert B., Elizabeth Jockusch, and Curtis McKnight.  
 "Cognitive Processes in Learning Algebra." *Journal of Children's Mathematical Behavior*, Vol. 2, #1 (Spring 1978), pp. 10-320.
- Davis, Robert B. (with Leon Henkin).  
 "Inadequately Tested Aspects of Mathematics Learning." Pt. II (Davis)  
 "Aspects of Mathematics Learning that Should Be the Subject of Testing." Paper presented at NIE Conference on Cognition, Technology, and Assessment, August, 1978.
- Davis, Robert B.  
 "Children's Learning of Mathematics," in *L'Insegnamento integrato delle scienze nella scuola primaria*. Rome: Accademia Nazionale Dei Lincei, 1979.
- Davis, Robert B.  
 "Conceptualizing the Structures Underlying Cognitive Behavior - the Usefulness of 'Frames'." Paper presented at American Educational Research Association Annual Meeting, San Francisco, April 1979.
- Davis, Robert B.  
 "Error Analysis in High School Mathematics, Conceived as Information-Processing Pathology." Paper presented at the American Educational Research Association Annual Meeting, San Francisco, April 1979.
- Davis, Robert B. and Curtis C. McKnight.  
 "Towards Eliminating 'Black Boxes': A New Look at Good vs. Poor Mathematics Students." Paper presented at the American Educational Research Association Annual Meeting, San Francisco, April 1979.
- Davis, Robert B. and Curtis C. McKnight.  
 "Modeling the Processes of Mathematical Thinking." *Journal of Children's Mathematical Behavior*, Vol. 2, no. 2 (spring 1979).
- McKnight, Curtis.  
 "The Use of 'Frames' in Conceptualizing Learning and Problem Solving in Mathematics." Paper presented at the A.E.R.A. Annual Meeting, San Francisco, April, 1979.
- McKnight, Curtis C.  
 "Perception of Algebraic Expressions: Features and Organization." Unpublished paper, 1978.
- Davis, Robert B., Curtis McKnight, Philip Parker, and Douglas Elrick.  
 "Analysis of Student Answers to Signed Number Arithmetic Problems." *Journal of Children's Mathematical Behavior*, Vol. 2, #2 (Spring, 1979).
- A very large number of requests for reports have been received by the project; in addition, some of the AERA presentations were made to standing-room-only audiences.
- As nearly as we can tell, the work seems to have been unusually well received by professional audiences of various sorts. We already have invitations for future presentations extending through August, 1980.

## Bibliography

- Averch, N.S., S. Carroll, T. Donaldson, N. Kiesling, and J. Pincus.  
*How Effective Is Schooling?* Santa Monica, Calif.: RAND Corporation,  
 R-956-PCSF/RC, 1972, pp. 100-125.
- Bissel, J.  
 "Planned Variation in Head Start and Follow Through." In: J.S.  
 Stanley, ed., *Compensatory Education for Children, Ages 2 to 8*.  
 Baltimore, Md.: John Hopkins University Press, 1973, pp. 63-108.
- Bobrow, Daniel G.  
 "Dimensions of Representation." In: D.G. Bobrow and Allan Collins, eds.,  
*Representation and Understanding*. New York: Academic Press, 1975.
- Bobrow, D.G. and Donald A. Norman.  
 "Some Principles of Memory Schemata." In: D.G. Bobrow and Allan Collins,  
 eds., *Representation and Understanding*. New York: Academic Press, 1975.
- Brown, John Seely.  
 "Towards a Theory of Semantics of Procedural Skills in Mathematics."  
 Paper presented at the Annual Meeting of the American Educational Research  
 Association, San Francisco, California, April, 1979.
- Brown, John Seely and Richard R. Burton.  
 "Diagnostic Models for Procedural Bugs in Basic Mathematical Skills."  
*Cognitive Science*, vol. 2, no. 2 (April-June, 1978).
- Callahan, Raymond E.  
*Education and the Cult of Efficiency*. Chicago: University of Chicago  
 Press, 1962.
- Carr, H. and J.B. Watson.  
 "Orientation in the White Rat." *Journal of Comparative Neurology and*  
*Psychology*, vol. 18, 1908, pp. 27-44.
- C.B.M.S. National Advisory Committee on Mathematical Education.  
*Overview and Analysis of School Mathematics, Grades K-12*. Washington, DC:  
 Conference Board of the Mathematical Sciences (2100 Pennsylvania Ave., N.W.,  
 Suite 832, Washington, DC 20037), 1975.
- Clement, John and James Kaput.  
 Letter to the Editors. *Journal of Children's Mathematical Behavior*,  
 vol. 2, no. 2 (spring, 1979).
- Cochran, Beryl, Alan Barson, and Robert B. Davis.  
 "Child-created Mathematics." *The Arithmetic Teacher*, vol. 17, no. 3 (March,  
 1970), pp. 211-215.
- Cremin, Lawrence.  
*The Transformation of the School*. New York: Vintage Books, 1961.

Davis, Robert B.

*Discovery in Mathematics: A Text for Teachers.* Addison-Wesley, 1964.

Davis, Robert B.

"The structure of Mathematics and the Structure of Cognitive Development." *Journal of Children's Mathematical Behavior*, vol. 1, no. 1 (winter, 1971-72), pp. 71-99.

Davis, Robert B.

"Representing Knowledge About Mathematics for Computer-Aided Teaching, Part I. --Educational Applications of Conceptualizations from Artificial Intelligence." In: E.W. Elcock and Donald Michie, eds., *Machine Intelligence 8: Machine Representations of Knowledge*. Ellis Horwood, Ltd. (dist. by Halstead Press, John Wiley, & Sons, Inc., New York), 1977, pp. 363-386.

Davis, Robert B.

"Conceptualizing the Structures Underlying Cognitive Behavior - the Usefulness of 'Frames.'" Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, California, April, 1979. [1979A]

Davis, Robert B.

"Error Analysis in High School Mathematics, Conceived as Information-Processing Pathology." Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, California, April, 1979. [1979B]

Davis, Robert B., Elizabeth Jockusch, and Curtis McKnight.

"Cognitive Processes in Learning Algebra." *Journal of Children's Mathematical Behavior*, vol. 2, no. 1 (spring, 1978), pp. 10-320.

Davis, Robert B. and Leon Henkin.

"Aspects of Mathematics Learning that Should Be the Subject of Testing." In: Ralph Tyler and Sheldon White, eds., *Testing, Teaching, and Learning: Invited Papers*. Washington, DC: National Institute of Education, 1979, pp. 26-48.

Davis, Robert B. and Curtis C. McKnight.

"Modeling the Processes of Mathematical Thinking." *Journal of Children's Mathematical Behavior*, vol. 2, no. 2 (spring, 1979).. [1979A]

Davis, Robert B., Curtis C. McKnight, Philip Parker, and Douglas Elrick.

"Analysis of Student Answers to Signed Number Arithmetic Problems." *Journal of Children's Mathematical Behavior*, vol. 2, no. 2 (spring, 1979).

Davis, Robert B., Tom Romberg, Sidney Rachlin, and Mary Grace Kantowski.

*Research and Development in Mathematics Education in the Soviet Union.* (in preparation)

D.H.E.W.

*The Effectiveness of Compensatory Education: Summary and Review of the Evidence.* Washington, D.C.: Dept. of Health, Education, and Welfare, April, 1972.

Dorcus, R.M. and W.L. Gray.

"The Role of Kinesthesia in Retention by Rats." *Journal of Comparative Psychology*, vol. 13, 1932, pp. 447-451.

Erlwanger, Stanley H.

"Benny's Conception of Rules and Answers in IPI Mathematics." *Journal of Children's Mathematical Behavior*, vol. 1, no. 2 (autumn, 1973), pp. 7-26.

Erlwanger, Stanley H.

*Case Studies of Children's Conceptions of Mathematics.* Doctoral Dissertation, University of Illinois, Urbana, Illinois, 1974.

Erlwanger, Stanley H.

"Case Studies of Children's Conceptions of Mathematics--Part I." *Journal of Children's Mathematical Behavior*, vol. 1, no. 3 (summer, 1975), pp. 157-283.

Feurzeig, Wallace & George Lukas.

"LOGO - A Programming Language for Teaching Mathematics." *Educational Technology*, March, 1972.

Friedman, Daniel P.

*The Little LISPer.* Chicago, IL: Science Research Associates, 1974.

Friend, Jamesine.

"Column Addition Skills." *Journal of Children's Mathematical Behavior*, vol. 2, no. 2 (spring, 1979), pp. 29-54.

Gagné, Robert M.

"The Acquisition of Knowledge." *Psychological Review*, vol. 59, 1962, pp. 355-365.

Gagné, Robert M.

"Curriculum Research and the Promotion of Learning." In: Ralph Tyler, Robert Gagné, and Michael Scriven, eds., *Perspectives of Curriculum Evaluation.* Chicago, IL: Rand McNally, 1967, pp. 19-38.

Goldstein, Ira.

"Will Computation Become the New Mathematics for Understanding Teaching and Learning Processes?" Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, California, April, 1979.

Gordon, E.W. and D.A. Wilkerson.

*Compensatory Education for the Disadvantaged.* New York: College Entrance Examination Board, 1966.

Hawkins, David.

"Learning the Unteachable." Chapter 1 in: Lee S. Shulman and Evan R. Keislar, eds., *Learning by Discovery: A Critical Appraisal.* Rand McNally, 1966, pp. 3-12.

- Heller, Joan I. and James G. Greeno.  
 "Information Processing Analyses of Mathematical Problem Solving." In:  
 Ralph Tyler and Sheldon White, eds, *Testing, Teaching, and Learning:  
 Invited Papers*. Washington, DC: National Institute of Education, 1979,  
 pp. 79-101.
- Henkin, Leon and Robert B. Davis.  
 "Inadequately Tested Aspects of Mathematics Learning." In: Ralph Tyler  
 and Sheldon White, eds., *Testing, Teaching, and Learning: Invited Papers*.  
 Washington, DC: National Institute of Education, 1979, pp. 49-63.
- Hilliard, Asa G.  
 "Standardized Testing and African-Americans: Building Assessor Competence  
 in Systematic Assessment." In: Ralph Tyler and Sheldon White, eds.,  
*Testing, Teaching, and Learning: Invited Papers*. Washington, DC: National  
 Institute of Education, 1979, pp. 170-184.
- Hinsley, Dan, John Hayes, and Herbert Simon.  
 "From Words to Equations -- Meaning and Representation in Algebra Word  
 Problems." In: Marcel Just and Patricia Carpenter, eds., *Cognitive  
 Processes in Comprehension*. Hillsdale, NJ: Lawrence Erlbaum, 1977.
- Hoffman, Banesh.  
*The Tyranny of Testing*. London: Crowell-Collier, 1962. (New York: Collier  
 Books, 1964)
- Houts, Paul L.  
*The Myth of Measurability*. New York: Hart Publishing Co., 1977.
- Kennedy, W.A., H. Willicutt, and A. Smith.  
 "Wechsler Profiles of Mathematically Gifted Adolescents." *Psychological  
 Reports*, vol. 12, 1963.
- Klahr, David.  
 "Designing a Learner." Chapter 19 in: David Klahr, ed., *Cognition and  
 Instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1976, pp. 325-331.
- Krutetskii, V.A.  
*The Psychology of Mathematical Abilities in Schoolchildren*. (Translated from  
 the Russian by Joan Teller. Edited by Jeremy Kilpatrick and Izaak Wirszup.)  
 Chicago, IL: University of Chicago, 1976.
- Kuhn, Thomas S.  
*The Structure of Scientific Revolutions*. Chicago: University of Chicago Press,  
 1962.
- Locke, E.A.  
 "Some Correlates of Classroom and Out-of-Class Achievement in Gifted Science  
 Students." *Journal of Educational Psychology*, vol. 54, 1963.
- Matz, Marilyn.  
 "Underlying Mechanisms of Bugs in Algebraic Solutions." Paper presented  
 at the Annual Meeting of the American Educational Research Association,  
 San Francisco, California, April, 1979.
- Mays, John M., Arthur S. Melmed, and Sylvia Scribner.  
 Preface to: Ralph Tyler and Sheldon White, eds., *Testing, Teaching and  
 Learning: Invited Papers*. Washington, DC: National Institute of Education, 1979.

Miller, George.

"The Magic Number  $7 \pm 2$ ." *Psychological Review*, vol. 63, 1956, pp. 81-97.

Miller, George.

Quoted in: Ralph Tyler and Sheldon White, *Testing, Teaching, and Learning*. Washington, DC: National Institute of Education, 1979, p. 1.

Minsky, Marvin.

"A Framework for Representing Knowledge." In: P. Winston, ed., *The Psychology of Computer Vision*. New York: McGraw-Hill, 1975.

Minsky, Marvin and Seymour Papert.

*Artificial Intelligence Memo No. 252*. Cambridge, MA: Massachusetts Institute of Technology, Artificial Intelligence Laboratory, January 1, 1972.

Morrison, Philip.

"The Bell Shaped Pitfall." In: Paul Houts, ed., *The Myth of Measurability*. New York: Hart Publishing Co., 1977, pp. 82-89.

Murray, J.E.

"An Analysis of Geometric Ability." *Journal of Educational Psychology*, vol. 40, 1949.

Newell, Allen, J.C. Shaw, and Herbert Simon.

"Elements of a Theory of Human Problem Solving." *Psychological Review*, vol. 65, no. 3 (1958), pp. 151-166.

Newell, Allen and Herbert A. Simon.

*Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall, 1972.

Olton, Davis S.

"Mazes, Maps, and Memory." *American Psychologist*, vol. 34, no. 7 (July, 1979), pp. 583-596.

Padilla, Amado.

"Critical Factors in the Testing of Hispanic Americans: A Review and Some Suggestions for the Future." In: Ralph Tyler and Sheldon White, eds., *Testing, Teaching, and Learning: Invited Papers*. Washington, DC: National Institute of Education, 1979, pp. 185-209.

Papert, Seymour.

"Computer-Based Microworlds as Incubators for Powerful Ideas." Cambridge, MA: Massachusetts Institute of Technology, Artificial Intelligence Laboratory, March 1, 1978.

Polanyi, Michael.

*Personal Knowledge*. New York: Harper and Row, 1958.

Resnick, Cecily.

*Computational Models of Learners for Computer Assisted Learning*. Unpublished Ph.D. thesis, University of Illinois, Urbana, Illinois, 1975.

Restle, Frank.

"Discrimination of Cues in Mazes: A Resolution of the 'Place-vs.-Response' Question." *Psychological Review*, vol. 64, no. 4 (July, 1957), pp. 217-228.

Rumelhart, David E. and Andrew Ortony.

"The Representation of Knowledge in Memory." Chapter 4 in: Richard C. Anderson, Rand Spiro, and William Montague, eds., *Schooling and the Acquisition of Knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

Schank, Roger C.

"The Structure of Episodes in Memory." In: D.G. Bobrow and Allan Collins, eds., *Representation and Understanding*. New York: Academic Press, 1975.

Schank, Roger.

"Representation and Understanding of Text." Chapter 27 in: E.W. Elcock and Donald Michie, eds., *Machine Intelligence 8: Machine Representations of Knowledge*. Ellis Horwood, Ltd. (dist. by Halstead Press, John Wiley and Sons, Inc., New York), 1977, pp. 575-619.

Schwartz, Judah L.

"A is to B is to C is to Anything At All: the Illogic of IQ Tests." In: Paul Houts, ed., *The Myth of Measurability*. New York: Hart Publishing Co., 1977, pp. 90-99. [1977A]

Schwartz, Judah L.

"Math Tests." In: Paul Houts, ed., *The Myth of Measurability*. New York: Hart Publishing Co., 1977, pp. 282-290. [1977B]

Schwartz, Judah L. and Edwin F. Taylor.

"Project TORQUE." In: Ralph Tyler and Sheldon White, eds., *Testing, Teaching, and Learning: Invited Papers*. Washington, DC: National Institute of Education, 1979, pp. 227-238.

Simon, Herbert.

Statement made at the Greeno Conference on Cognitive Aspects of the Learning of Algebra, Pittsburgh, PA, July 30-31, 1979.

Simon, Herbert A.

"Cognitive Processes and the Structure of Knowledge in Physics and Algebra." Research now in progress at Carnegie-Mellon University, Pittsburgh, PA.

Small, W.S.

"Experimental Study of the Mental Processes of the Rat. II." *American Journal of Psychology*, vol. 12, 1901, pp. 206-239.

Stanley, Julian C., ed.

*Compensatory Education for Children, Ages 2 to 8*. Baltimore, Md.: Johns Hopkins University Press, 1973.

Stein, Sherman K. and Calvin D. Crabill.

*Elementary Algebra: A Guided Inquiry*. Boston, MA: Houghton-Mifflin, 1972.

Taylor, Edwin F.

"Science Tests." In: Paul Houts, ed., *The Myth of Measurability*. New York: Hart Publishing Co., 1977, pp. 291-308.

Tyler, Ralph W. and Sheldon H. White. *Testing, Teaching, and Learning*.

Washington, DC: National Institute of Education, 1979 [1979A].

- Tyler, Ralph W. and Sheldon H. White, eds.  
*Testing, Teaching, and Learning: Invited Papers.* Washington, DC: National Institute of Education, 1979 [1979B].
- Van Lehn, Kurt and John Seely Brown.  
 "Planning Nets: A Representation for Formalizing Analogies and Semantic Models of Procedural Skills." In: R.E. Snow, P.A. Frederico, and W.E. Montague, eds., *Aptitude Learning and Instruction: Cognitive Process Analyses.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.
- Weber, George.  
 "The Cult of Individualized Instruction." *Bulletin of the Council for Basic Education*, vol. 20, no. 9 (May, 1976), p. 104.
- Winograd, Terry.  
*Procedures as a Representation for Data in a Computer Program for Understanding Natural Language.* Cambridge, MA: Massachusetts Institute of Technology, Artificial Intelligence Laboratory, Feb. 1971.
- Winston, Patrick H.  
 "Learning by Creatifying Transfer Frames." *Artificial Intelligence*, vol. 10, no. 2 (April, 1978), pp. 147-172.
- Zacharias, Jerrold R.  
 "The Trouble With IQ Tests." In: Paul Houts, ed., *The Myth of Measurability.* New York: Hart Publishing Co., 1977, pp. 66-81.
- Zacharias, Jerrold R.  
 "A Discussion of Tests that Concern the Sciences and the Technologies." In: Ralph Tyler and Sheldon White, eds., *Testing, Teaching, and Learning: Invited Papers.* Washington, DC: National Institute of Education, 1979, pp. 20-25.
- Zankov, L.V.  
*Teaching and Development: A Soviet Investigation.* White Plains, NY: M.E. Sharpe, 1977.